DOCUMENT RESUME

ED 283 027 CE 047 499

AUTHOR	Sipe, Cynthia L.; And Others
TITLE	Summer Training and Education Program (STEP). Report on the 1986 Experience.
INSTITUTION	Public/Private Ventures, Philadelphia, PA.
PUB DATE	Apr 87
NOTE	283p.; For the 1985 report, see CE 047 500.
AVAILABLE FROM	Public/Private Ventures, 399 Market Street, Philadelphia, PA 19106.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Attitude Change; *Basic Skills; Comparative Analysis; Demonstration Programs; *Disadvantaged Youth; *Economically Disadvantaged; Geographic Location; National Programs; Outcomes of Education; Pretests Posttests; Program Effectiveness; *Remedial Instruction; Secondary Education; *Summer Programs; *Youth Programs
IDENTIFIERS	Impact Studies; *Summer Training and Education Program

ABSTRACT

In the summer of 1986, 564 of the 752 members of the treatment group of the first full-scale implementation of the Summer Training and Education Program (STEP) model returned for their second summer. At the same time, a second cohort consisting of 1,591 youth (765 in the treatment group and 754 in the control group) began their first summer of participation. STEP had substantively large effects on the Cohort II participants' reading and math skills, knowledge of contraception, and sexually related behavior during their first summer of participation. The effects of program participation on basic skills were larger and much more widespread than those for Cohort I's first summer (1985) and thus are more likely to result in the desired longer-term outcomes. The Cohort II control group youths experienced substantial learning losses in both reading and math over the 1986 summer, as had the Cohort I control group youth in the summer of 1985. The research also examined STEP's impact on the school-year (1985-1986) and second-summer (1986) performance of Cohort I youths. Those in the Cohort I treatment group were 22 percent less likely to fail a grade than were their counterparts in the control group. The impact was especially large for Hispanics. Other school-year effects from participation in STEP appeared to be small. (Appendixes include a program participation analysis, a description of the Summer Work Enhancement Program, technical data tables, tables referenced in the text, and figures on the total cost of STEP in 1986.) (MN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED283027

Public/Private Ventures

399 Market Street
Philadelphia, PA 19106
(215) 592-9099

Summer Training and Education Program (STEP)

Report on the 1986 Experience

by Cynthia L. Sipe
Jean Baldwin Grossman
Julita A. Milliner

CE047499

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

April 1987

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY
A Jaffe
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

ACKNOWLEDGMENTS

The report on the 1986 Summer Experience of STEP was completed through the contributions of many individuals. The authors wish to acknowledge the support provided and extend their sincere gratitude to all of the report's contributors. First, we would like to thank members of the P/PV staff: Wendy Wolf, Director of Research at P/PV, supervised the production of the report and provided invaluable insight and advice on research design and analytical issues; the data unit staff--John Wooden, Rick Etheredge, Ken Libert and Jon Bumbaugh--provided programming and analytic assistance, while Batia Trietsch, Eleanor Hammond and Kevin Harris assisted with data management; and Mike Bailin, Gary Walker, Jerry Kolker, Frances Vilella-Velez, Mike Sack and May Pritchard reviewed and commented on early drafts of the report. Editorial support was provided by Natalie Jaffe, Judy Wilmoth White and Cary Lawson; and Paula Strawberry typed the report. Finally, we wish to acknowledge the contributions of P/PV's site coordinators and their assistants in collecting and providing relevant data.

Several individuals outside of P/PV also contributed to the report. We wish to acknowledge the contributions made by David Zimmerman and Bob Egan in collecting and analyzing the cost data for the program. Calvin Jones, Frank Furstenberg, Walter Haney, Barbara Heyns and Don Rock provided advice on general research issues. Carolyn Bumbaugh designed the figures. Gordon Berlin, Eileen Rudert and numerous others read and provided useful comments on an earlier draft of the report.

CONTENTS

Page

EXECUTIVE SUMMARY i

I. INTRODUCTION 1

BACKGROUND 2
POLICY CONTEXT 6
ORGANIZATION OF THIS REPORT 8

II. THE STEP MODEL AND EARLY FINDINGS 9

PROGRAM COMPONENTS 9
TARGET POPULATION 12
THE PILOT EXPERIENCE 12
SITE SELECTION AND DESCRIPTION 13
THE 1985 SUMMER EXPERIENCE 15

III. RESEARCH DESIGN 17

IMPLEMENTATION ANALYSIS 17
IMPACT ANALYSIS 18
DESIGN STRATEGY 19
DATA COLLECTION STRATEGY 20
ANALYTIC STRATEGIES 26

IV. IMPLEMENTING STEP IN THE SECOND SUMMER 31

PARTICIPANT SELECTION 31
TESTING 36
THE REMEDIATION COMPONENT 40
THE LIFE SKILLS AND OPPORTUNITIES (LSO) COMPONENT 49
THE WORK EXPERIENCE COMPONENT 55
THE SCHOOL-YEAR SUPPORT (SYS) COMPONENT 57
SUMMARY 59

V. PARTICIPANT BASELINE CHARACTERISTICS 61

THE COHORT II SAMPLE 61
THE COHORT I RETURNEES 74
SUMMARY OF PARTICIPANT CHARACTERISTICS 83

VI. FIRST SUMMER IMPACTS--COHORT II AND AGGREGATE SAMPLES 85

ACADEMIC PERFORMANCE 85
LIFE SKILLS OUTCOMES 106

CONTENTS - (cont'd)

Page

VII. SCHOOL YEAR AND SECOND SUMMER IMPACTS FOR COHORT I 117

SCHOOL DATA ANALYSIS 117
CHARACTERISTICS OF COHORT I RETURNEES 130
DESCRIPTION OF SECOND SUMMER AND 15-MONTH OUTCOMES 133

VIII. COST ANALYSIS 147

METHODOLOGY 147
THE COSTS OF STEP 153
TOTAL 15-MONTH PROGRAM COSTS 159

IX. SUMMARY AND CONCLUSIONS 163

STEP IMPLEMENTATION DURING THE SECOND SUMMER 163
PARTICIPANTS 165
SUMMER IMPACTS 165
SCHOOL-YEAR IMPACTS FOR COHORT I 167
SECOND SUMMER RESULTS FOR COHORT I 167
COSTS 168
IMPLICATIONS 168

REFERENCES 171

APPENDICES

A. Program Participation Analysis 175

B. Summer Work Enhancement Program (SWEP)....179

C. Tables Referenced in Text 185

D. Total Cost of STEP in 1986 211

E. Technical Data Tables 213

TABLES

Page

III.1 Summer Training and Education Program Data Base by Source 21

IV.1 Academic Eligibility Measures by Site 32

IV.2 Re-enrollment of Cohort I Youth in 1986 37

IV.3 Youth Tested by Site for Both Cohorts 39

IV.4 Sample Weekly Summer Program Schedule 42

IV.5 Mean Hours (Percent) of Participation in Classroom and Work Experience Components 50

V.1 Selected Characteristics of the Randomization and Learning Analysis Samples for Cohort II 63

V.2 Demographic and Educational Ability Characteristics of Cohort II..................... 65

V.3 Additional Demographic and Economic Characteristics of Cohort II 68

V.4 School Participation and Educational Expectations Characteristics of Cohort II 70

V.5 Fertility-related Characteristics of Cohort II .. 72

V.6 Demographic and Educational Ability Characteristics of the Cohort I Returning Treatments 76

V.7 Additional Demographic and Economic Characteristics for Cohort I Returning Treatments 77

V.8 School Participation Characteristics and Educational Expectations of Cohort I Returning Treatments 79

V.9 Fertility-related Characteristics of Cohort I Returning Treatments 81

VI.1 Mean Control Group Losses in Reading and Math in Grade Equivalents by Site for the Pilot Sample, Cohort I and Cohort II 87

TABLES - (cont'd)

Page

VI.2 Means of Pre-test to Post-test Changes in Reading and Math Scores by Sex and Race/Ethnicity for Cohort II 92

VI.3 Means of Pre-test to Post-test Changes in Reading and Math Scores by Site for Cohort II 93

VI.4 Net Impact of Participation in STEP on MAT Scaled Scores in Reading and Math for Cohort II and Aggregate STEP Samples 95

VI.5 Means of Pre-test to Post-test Changes in Reading and Math Scores by Remediation Subsite in Boston, San Diego and Seattle for Cohort II 99

VI.6 Cohort II Youth Who Re-enrolled in School, Fall 1986, Overall and by Sex, Race/Ethnicity and Site 104

VI.7 Net Impact of Participation in STEP on Birth Control Knowledge Variables for Cohort II and Aggregate by Sex, Race/Ethnicity and Site 108

VI.8 Baseline to Endline Changes in Percentage of Cohort II Youth with Sexual Experience Having Sex in the Previous Two Months and Using Contraceptives at Last Intercourse 112

VII.1 The Estimated Impacts of STEP on Absenteeism, and Test Scores in Reading and Math During the Intervening School Year by Sex, Race/Ethnicity and Site for Cohort I 124

VII.2 The Estimated Impacts of STEP on Scaled Test Scores During the Intervening School Year by Site for Cohort I 126

VII.3 The Estimated Impacts of STEP on Credits Earned and Grade Promotion During the Intervening School Year by Sex, Race/Ethnicity and Site for Cohort I...129

VII.4 Selected Characteristics of Returning Cohort I Treatment Youth and San Diego Control Youth at Beginning of Second Summer, by Site 132

VII.5 Changes in Mean Reading and Math Scores Over 15 Months for Returning Cohort I Treatments and San Diego Controls 135

TABLES - (cont'd)

Page

VII.6 Mean Changes in Knowledge of Contraceptive Methods and Availability Over 15 Months for Returning Cohort I Treatments and San Diego Controls 141

VII.7 Baseline to Endline Changes in Mean Health Knowledge and Mean Drug Knowledge for Cohort I ..145

VIII.1 The Synthetic Cohort Approach: Sources of Component Costs 149

VIII.2 Estimated Costs Per Cohort Member by Site for 15 Months of STEP 154

FIGURES

VI.1 Mean MAT Reading Scores for Cohort II At Pre- and Post-test 90

VI.2 Mean MAT Math Scores for Cohort II at Pre- and Post-test 90

VI.3 Pre-Test and Post-Test Changes in Mean Reading Scores for Cohort II by Sex and Race/Ethnicity ...91

VI.4 Pre-Test to Post-Test Changes in Mean MAT Math Scores for Cohort II by Sex and Race/Ethnicity....91

EXECUTIVE SUMMARY

The need to identify effective ways of improving the basic skills of disadvantaged youth has become an increasingly urgent and widespread concern throughout American society.

This concern has been accentuated by the impact of major demographic changes that are taking place: while youth are becoming a smaller segment of our total population, the number of youth in poor families is growing--about one in five teenagers is now living in poverty. Since the educational achievement of poor youth is on average lower than that of more advantaged youth, the basic skills competence of the entry-level work force is on the decline.

Yet changes in public education inspired by the reform movement of the early 1980s have done little to improve the performance or retention of youth who are doing poorly in school or dropping out. Those reforms have, rather, focused on upgrading the achievement of youth already performing acceptably. A growing share of the youth population thus seems to be falling behind relative to the basic skills requirements of an increasingly complex society and economy.

But while the search for effective solutions is urgent, the nation's concern about the budget deficit has led to reduced federal expenditures on social programs and has largely restricted interventions to those that can be implemented at moderate cost. The Summer Training and Education Program (STEP) is such an intervention.

The STEP model aims to increase basic skills and lower dropout and teen pregnancy rates by providing poor and under-performing youth with remediation, life skills and work experience during two consecutive and intensive summer programs, with ongoing support and personal contact during the intervening school year. It builds on and enriches existing public services: work experience provided by the federal Summer Youth Employment and Training Program (SYETP), and education provided by public school resources. Thus, the model requires only moderate additional expenditures to implement.

STEP was designed and initiated by Public/Private Ventures in 1984. The model is being tested in a five-site national demonstration that includes a four-year operational phase and research activities that extend for an additional five years.

Initial funding for model development and pilot testing, and continuing support for the national demonstration have been provided by The Ford Foundation. Since the national demonstration began in the summer of 1985, support has expanded to include

Aetna Life and Casualty Foundation, The Ahmanson Foundation, The Edna McConnell Clark Foundation, William T. Grant Foundation, The William and Flora Hewlett Foundation, The Robert Wood Johnson Foundation, Lilly Endowment, James C. Penney Foundation, the U.S. Department of Health and Human Services, and the U.S. Department of Labor. Local program operation costs are covered primarily by Title II-B funds under the Job Training Partnership Act (JTPA) and local school district resources.

STEP's early impacts are very encouraging. In the 1986 summer program, new enrollees largely held their own in reading while their control group counterparts showed substantial losses. Likewise, treatment youth achieved gains in math while the control group lost ground. Stemming summer learning losses--which research has shown poor youth experience to a far greater degree than their more advantaged peers--is an important first step in keeping youth in school and improving their performance.

The summer program's life skills component also had encouraging results. STEP youth increased their knowledge of the consequences of teen parenting and how to avoid it, and were 53 percent more likely than the control group to use contraceptives if they were sexually active (nearly half the youth reported at the beginning of the summer that they were).

THE SUMMER TRAINING AND EDUCATION PROGRAM (STEP) DEMONSTRATION

The STEP demonstration was designed to respond to the following circumstances:

- Low-income youth are dropping out of high school at alarming rates -- 50 percent and higher in many urban areas.
- The relationship between dropping out of school and long-term difficulty in the labor market is well documented, as is the inter-relationship of basic skills deficiencies, teenage parenting and dropping out.
- The early teen years are particularly critical for dropout-prone youth; their school experience at that time strongly influences later decisions to drop out or to graduate.
- Economically disadvantaged youth lose more ground academically during the summer months than do their more advantaged peers.

- The highest priority for most poor youth during the summer is securing income; thus, since jobs are scarce for 14 and 15 year-olds, a summer program with a paycheck is highly attractive.

Recognizing the relationship of these factors--and the need for interventions that are both operationally and economically feasible--the STEP model comprises a two-summer program during which participants earn minimum wage for a full day, five days a week for six-eight weeks, and engage in three core activities:

1. Remediation: 90 hours of group and individually paced instruction in basic reading and math skills.

   About 20 percent of this time is spent on computer-assisted instruction, and 20 minutes a day are spent on silent sustained reading. The curriculum, called Practical Academics, was developed by P/PV and includes nine modules that teach basic skills in contexts relevant to these students--developing job skills, life skills and the ability to profit from regular school classes. The modules can be adjusted to meet local needs.

2. Life Skills and Opportunities (LSO): 18 hours of instruction on responsible social and sexual attitudes and behavior.

   Emphasis is on personal decision-making; job equality issues; the consequences of sexual activity, teen pregnancy and substance abuse; ways to avoid pregnancy, including abstinence; and sources of family planning services. The curriculum, developed by P/PV with outside consultants, uses lectures, discussions, films, role-playing, field trips and outside speakers to stress the need to set goals, plan for the future and take responsibility for a decision about whether or not to initiate sexual activity. About half the youth report on pre-program questionnaires that they are sexually active but not knowledgeable about contraception.

3. Work experience: at least 80 hours of part-time work provided by the federally funded Summer Youth Employment and Training Program (SYETP).

   Combining SYETP jobs with remediation and life skills instruction and paying the minimum wage for all a participant's hours provide financial incentives to youth who ordinarily might not become involved in academic remediation and pregnancy prevention programs. It also demonstrates a way to institutionalize an enriched summer youth jobs program.

During the intervening school year, STEP offers activities designed to encourage youth to remain in school. Youth participate in group activities and meet regularly with mentors/counselors who refer them to needed services, monitor their school attendance and encourage them to maintain the progress made in the first summer and return for the second.

To be eligible for participation, youth must be 14 or 15 years old, from low-income families and performing below grade level in reading or math as indicated by recent standardized test scores and/or a recent history of grade retention. Such youth are at high risk of dropping out of school.

The demonstration includes a research component that involves random assignment of youth to treatment and control groups. Control group youth work full time on SYETP jobs. Both groups are tested at the beginning and end of the summer, using the Metropolitan Achievement Test (MAT). Data collected from the school districts in each site include attendance, standardized test scores, credits earned, grade progression and dropout status. Pre- and post-program questionnaires assess attitudes, knowledge and behavior with respect to sexual and social issues and career awareness. After the operational period, follow-up research will continue through the scheduled graduation date of all youth; data will be collected from the schools and from the youth themselves through face-to-face and telephone interviews.

For each cohort, approximately 300 youth are recruited in each site; by random assignment, half are selected for treatment and half for a control group. In the second summer, treatment youth are encouraged to return to continue the enriched summer program; control group youth are not guaranteed summer jobs in most sites. (In the summer of 1986, only one site re-enrolled control youth in SYETP.)

The five demonstration cities are Boston, Fresno, Portland (Oregon), San Diego and Seattle. At the sites, day-to-day management and operation of the program are the responsibilities of employment and training agencies, school districts and a number of other local institutions. In all sites, public school involvement with recruitment, curriculum design and the provision of data for this summer program has been extensive, more so than in most employment and training programs. In addition, all sites grant school credit to students who participate in STEP remediation.

As initially designed, the STEP demonstration was to start in 1985, serve 1,500 youth in two overlapping waves (1985-86, Cohort I; 1986-87, Cohort II) and be followed by a long-term research phase (1987-92). Based on the success of the 1986 summer of program operations and the strength of its test and other results, however, the demonstration has been expanded to serve a third cohort of youth in 1987 and 1988, extending the follow-up phase until 1993. As a result, the total number of treatment group youth in the demonstration will be about 2,250; an equal number will compose the control group. (See Figure 1.)

THE 1985 SUMMER

Summer program activities for Cohort I began in 1985. The program was successful in reducing by half the substantial learning losses that would have occurred over the summer without the program. Treatment youth outscored their control group counterparts in both reading and math by approximately one-quarter of a grade equivalent. The program was not powerful enough, however, to cancel the summer losses for treatment group members. At the end of the summer, their test scores in both reading and math were lower than they were at the beginning of the summer, but control group losses were significantly larger than those experienced by treatment youth.

STEP also had a substantial effect on participants' knowledge of contraception. Information about birth control methods and availability significantly increased for every site, racial/ethnic group and sex.

These results were both encouraging and indicative of the areas in which the program needed to be strengthened. As a result, management and structure of the remediation component was improved and the curriculum was revised for the summer of 1986. The experience of Cohort I also indicated the need for a more intense and structured school-year component. P/PV strengthened the design, and sites introduced it at the beginning of the 1986-87 school year.

THE 1986 SUMMER

In the summer of 1986, 564 of the 752 treatment youth in Cohort I returned for their second summer of participation. A second cohort of 1,519 youth (765 treatments and 754 controls) began their first summer of program participation.

Analysis of the characteristics of the two cohorts indicates that they are similar in most respects: mostly minority, poor and from female-headed households; testing about two grades below the level appropriate for their age in reading and about one and one-half grades below in math; sexually active and largely uninformed about birth control methods. The two cohorts differ slightly in two areas: 55 percent of Cohort II is 14 years old, 45 percent is 15, while Cohort I was evenly split; and there are larger percentages of Hispanics and Asians in Cohort II.

As for the Cohort I treatment youth who returned for their second summer, they were somewhat more likely than non-returnees to be

SUMMER TRAINING AND EDUCATION PROGRAM (STEP)
OPERATIONAL PLAN

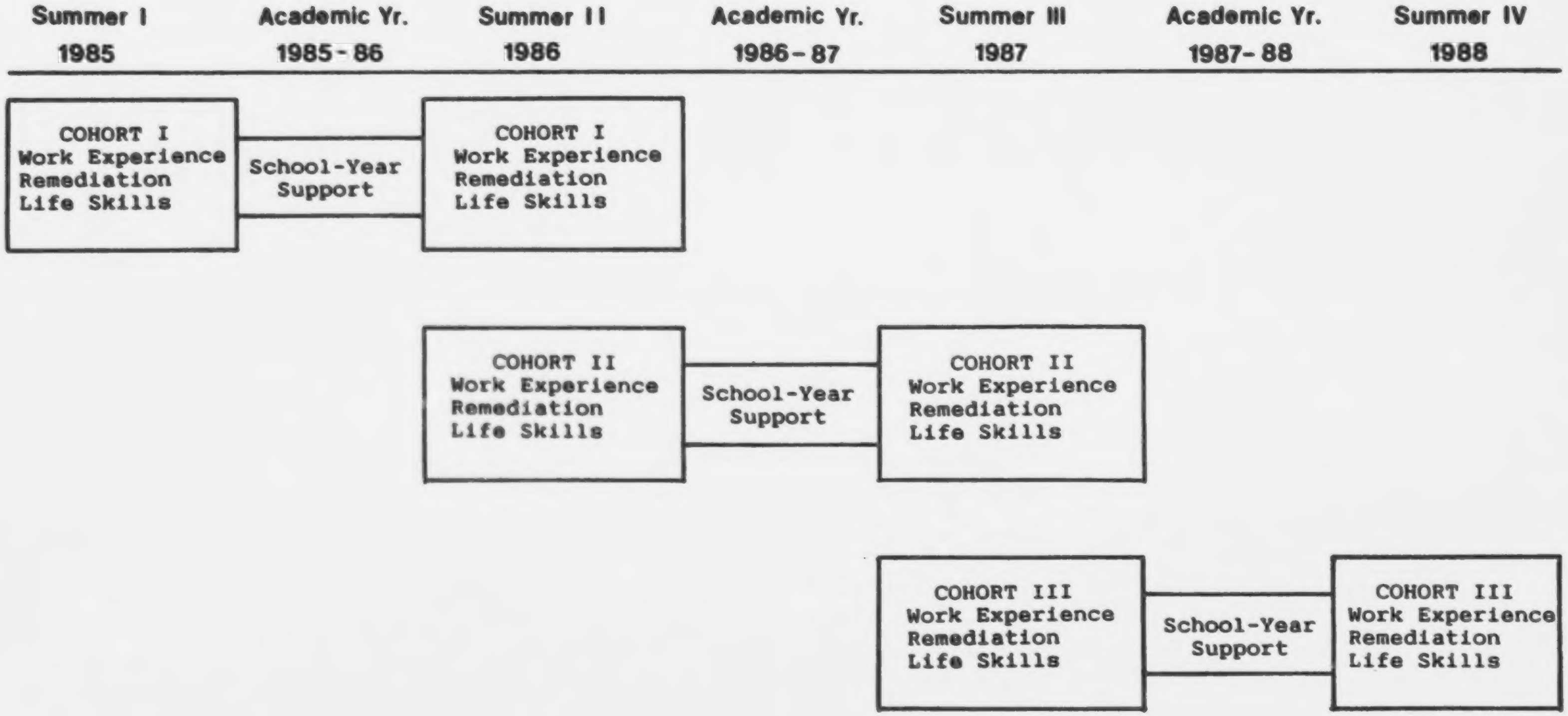

Figure 1

in the lower grades, to be from minority groups, and to have had no prior sexual or work experience.

Program participation was high among both cohorts. On average, treatment group youth were present for about 82 percent of their classroom hours and 80 percent of their work hours. As in Year One, the best predictor of participation levels was school-year attendance; youth who attended school regularly during the year continued that pattern during the summer.

1986 SUMMER IMPACTS

The Second Cohort

STEP had substantively large effects on Cohort II youths' reading and math skills, knowledge of contraception and sexually-related behavior during their first summer of participation. The effects on basic skills were larger and much more widespread than those for Cohort I's first summer (1985) and thus are more likely to result in the desired longer-term outcomes.

Cohort II control group youth experienced substantial learning losses in both reading and math over the 1986 summer, as had Cohort I control youth in the summer of 1985. These losses ranged from about three-quarters of a grade to a full grade equivalent in reading, and about one-half of a grade equivalent in math.

STEP's impact on these losses during the 1986 summer was more than double the effect it had in the summer of 1985. The majority of learning loss in reading was stemmed: STEP youth scored six-tenths of a grade equivalent higher than control youth. Learning loss in math was not only eliminated; a gain was produced. At the end of the summer, treatment youth scored slightly higher in math than they had at the beginning of the summer, and eight-tenths of a grade equivalent higher than controls.

These results were not only stronger but were also more consistent for the whole Cohort II treatment sample than they were for Cohort I's first summer. Treatment youth of both sexes, in all racial/ethnic groups and in all five demonstration sites substantially improved in math; in Cohort I's first summer only Hispanics, blacks and treatment youth in two of the five sites improved in math. In reading, both sexes, black youth and youth in four of the demonstration sites benefited; in Cohort I's first summer only females, Hispanics, Asians and treatment youth in two sites benefited.

The sharp improvement in basic skills likely reflects the improvements in design and operation of STEP remediation in 1986. P/PV developed and delivered to the sites a stronger, more prescriptive curriculum; P/PV and the sites identified and hired lead teachers to supervise curriculum development and delivery; and the sites initiated local curriculum planning and began teacher training earlier in the spring than was possible the first year. In addition, the start-up problems endemic to the first year of any demonstration had been reduced: the revised remediation curriculum was the only element untried in the sites.

STEP's effects on the second cohort's knowledge, attitudes and behavior regarding sexual activity and consequences were similarly strong. Treatment youths' knowledge at the end of the summer, as measured by questions about the kinds and availability of contraceptives, increased 50 percent. These effects were found in all demonstration sites for both sexes and all racial/ethnic groups. With the exception of whites, youth in all subgroups also increased their knowledge of the burdens of adolescent pregnancy.

STEP also had some positive effects on sexual behavior: STEP participants were 53 percent more likely than youth in the control group to use contraceptives if they were sexually active during the summer.

The First Cohort

The research also examined STEP's impact on the school-year (1985-86) and second-summer (1986) performance of Cohort I youth.

To measure school-year performance, P/PV examined attendance, standardized test scores, credits earned, grade promotion and dropout behavior (approximately 95 percent of treatment youth returned to school after their first summer in STEP). The experience of treatment and control youth differed on only one of these measures--promotion. But the impact on promotion was large: treatment group youth were 22 percent less likely to fail than were controls: the failure rate, 24.0 percent for controls, was reduced to 18.7 percent for treatment youth. The positive effect was especially large for Hispanics.

The otherwise small effects on school performance reflect STEP's modest impact on Cohort I youth during their first summer in the program. After the 1985 summer, treatment youth had scored only a quarter of a grade equivalent higher than control youth in reading and math, and both groups had suffered significant losses. The small school-year effects may also reflect the absence of a fully operational school-year support component in every site during the year 1985-86.

The net impact of the program cannot be determined because control group performance was not measured during the second year, except in San Diego. During the summer of 1986, Cohort I treatment youth in all sites showed, again, significant losses in reading and little change in math. However, over the whole 15-month program period, these youth experienced a small, non-significant loss in reading and a substantial gain in math. Participation in the second summer continued to improve their knowledge of contraceptives and added an increased understanding of substance abuse as well.

How Cohort I STEP participants compare with similar youth not enrolled in STEP could be measured only in one site. In San Diego, Cohort I control youth were given a second summer of work experience in exchange for being tested. Measurement of this group's experience indicates that STEP was successful in stemming the treatment group's summer learning losses in that site. In spite of nearly equivalent first-summer losses and school-year gains among San Diego treatment and control youth, control youth there lost significantly more than treatment youth in both reading and math during the second summer. San Diego treatment youth also gained significantly more than control youth in knowledge about contraception and substance abuse.

COSTS

The average cost of providing one summer of work experience alone in the federal summer jobs program (SYETP) is approximately $700 per enrollee. The cost to society of providing one summer of STEP remediation and life skills instruction totals an additional $717 per enrollee, an increment that includes the value of contributed items, such as classroom space and computers. Few demonstration sites, in fact, had to pay for these items. The incremental cost is for a program in which operations have matured beyond start-up and 150 new enrollees are served each year.

For the whole 15-month treatment, which includes two summers and an active school-year support component, STEP's incremental cost was $1,600 per enrollee.

IMPLICATIONS FOR POLICY AND PRACTICE

The STEP experience and research to date provide some important insights into the need, viability and usefulness of summer-based strategies to improve basic skills and keep youth in school.

STEP's corroboration and extension of previous research on summer learning loss indicate the need. The size and consistency of losses in reading and math experienced by control group youth in both cohorts is startling. The magnitude of these summer learning losses seems disproportionate to the brief period of time that elapsed while youth were working on SYETP jobs.

Though there is still a great deal to be learned about the learning loss phenomenon--whether or not it is sensitive to the type of test administered, and if and how summer losses can be retrieved during successive school years--it calls for a re-assessment of the educational importance of the summer period.

Such a reassessment would focus on the summer's emptiness of content vis a vis basic skills, rather than on its brevity as an opportunity for intervention. STEP, in fact, lasted only seven weeks--with remediation offered for about five half-days a week--and yet it was sufficient to stem substantial losses in reading and produce actual gains in math.

The nationwide summer jobs program provides one opportunity to test alternatives. Starting this summer, localities are required to assess the basic skills of the 650,000 youth enrolled in the program and to provide remedial assistance to those with weak basic skills. A survey by the National Job Training Partnership found that the vast majority of localities are organizing or planning to organize a remediation component along with their summer work experience programs in 1987.

The STEP experience in remediation--particularly the differences in approach and test results between the 1985 and 1986 summers--should offer some useful insight into the level of resources and educational structures necessary to produce short-term test results that hold promise for long-term improvements in basic skills.

STEP's operational experience and test results seem to confirm both the feasibility and importance of extended educational programming for high-risk students. STEP differs from simple school-year extension, however, by integrating into the summer's academic instruction an opportunity for low-income youth to work and earn a salary, and to discuss and learn about key life options and their implications. These additional elements may be crucial in providing the economic incentive and practical knowledge necessary for continued participation in regular schooling.

Finally, the STEP experience to date demonstrates the feasibility of public education and employment/training institutions working together to provide innovative, multi-dimensional and effective programming for high-risk youth.

I. INTRODUCTION

The Summer Training and Education Program (STEP), an intervention designed to reduce dropout levels among poor and educationally deficient youth, focuses on two factors that are closely associated with dropping out: poor academic performance and adolescent parenthood. Through a comprehensive mix of work experience, basic skills remediation, and life skills and opportunities instruction, STEP seeks to improve reading and math skills, increase the graduation rate and reduce the incidence of teen parenthood among disadvantaged youth who are doing poorly in school.

STEP is a 15-month program comprising two consecutive summers of combined remediation, life skills instruction and work experience with various low-intensity services (such as mentors and group meetings) offered during the intervening school year. Following a successful pilot in three cities during the summer of 1984, a four-year national demonstration of STEP began in five cities--Boston, Massachusetts; Fresno, California; Portland, Oregon; San Diego, California; and Seattle, Washington--during the summer of 1985. To evaluate the impacts of the program, half the eligible youth interested in participating were randomly assigned to the treatment group and were offered STEP program services, and half were assigned to the control group and were offered a summer job under the federal Summer Youth Employment and Training Program (SYETP).

At the sites, day-to-day management and operation of the program are the responsibility of employment and training agencies, school districts and a number of other local institutions. Overall management of the demonstration is conducted by Public/Private Ventures (P/PV) of Philadelphia, a private not-for-profit corporation that seeks effective approaches to job training and education for the disadvantaged. Funding for the demonstration is provided by Aetna Life and Casualty Foundation, The Ahmanson Foundation, The Edna McConnell Clark Foundation, The Ford Foundation, W. T. Grant Foundation, The William and Flora Hewlett Foundation, The Robert Wood Johnson Foundation, Lilly Endowment, James C. Penney Foundation, the U.S. Department of Health and Human Services and the U.S. Department of Labor. Local program operation costs are partly absorbed by these funders, but they are primarily covered by use of Title II-B funds under the Job Training Partnership Act, local school district resources and assorted other sources.

Three cohorts of disadvantaged 14- and 15-year-old youth are participating in the demonstration in each of the five sites. The first cohort of youth entered STEP in the summer of 1985. Cohort I participants were eligible to receive a summer of instruction and work experience in 1985, to participate in STEP activities during the 1985-86 school year and to return for a second summer of instruction and work experience in the summer of 1986. The second cohort of youth entered the program in the summer of 1986. Cohort II participants were similarly eligible to participate in the summer components of STEP in 1986, to join in school-year support activities during the 1986-87 school year and to receive a second summer of instruction and work experience in the summer of 1987. The third cohort enters the program in the summer of 1987.

This report discusses the implementation of the demonstration in the second summer (1986) and school-year support activities during the 1985-86 school year, describes characteristics of both the new group of participants (Cohort II) and of the returning Cohort I youth and presents an analysis of the costs to operate STEP and the estimated impacts of participation in one and two summers of the STEP program. In particular, the effect of STEP on changes in reading and math skills over the summer, in the knowledge of contraceptive methods, in understanding the consequences of early parenting and in school-year performance is examined. All of these are short-term outcomes of STEP participation. Long-term outcomes, including graduation and early labor force participation rates, will be reported as data are collected and analyzed over the next five years.

The remainder of this chapter provides background on the problems of dropping out of high school and of teen pregnancy, and the policy issues relevant to the demonstration.

BACKGROUND

Every year large numbers of minority youth lacking high school degrees and adequate literacy skills enter the labor market. A recent assessment of literacy among young adults found that dropouts and minorities are disproportionately represented among individuals with low levels of literacy (Kirsch and Jungeblut, 1986). Their lack of basic skills impedes their chances of finding initial employment and limits their long-term employment and earnings prospects.

Increasing the educational attainment of youth at risk of dropping out of school increases their employment opportunities; this, in turn, improves their ability to be economically self-sufficient. Numerous studies find that additional years of schooling result in additional income and that, in particular, high school completion is an important determinant of future income. High school dropouts can expect lifetime earnings equal to only two-thirds of the earnings of high school graduates (U.S. Department of Commerce, 1983).

The gaps in earnings and employment between high school graduates and dropouts are not only substantial but growing. Between 1965 and 1979, the disparity in earnings grew by over 20 percent (Sum et al., 1983). During roughly the same period, the percentage of dropouts who found employment declined significantly (from 52.5% to 40% employed). Among minority groups, particularly blacks and Hispanics, the problem is acute. These groups are most affected by poverty and most likely to exhibit educational deficiencies that will limit their long-term chances to become economically independent and productive citizens.

The issue of the high school dropout rate and its implications will become even more important in the future. Demographic projections indicate that the proportion of the youth population that is disadvantaged and from minority groups is growing. These are exactly the groups with the highest dropout rates. The National Center for Education Statistics (1985), using data from the March 1982 Current Population Survey, estimates that among 20- to 24-year-olds, approximately 23 percent of black and 40 percent of Hispanic youth have not completed four years of high school. As the number of disadvantaged youth increases, the consequences of under-education and low levels of literacy will more greatly affect the larger society as well as the individuals themselves.

Over the next 10 to 15 years, the educational skills requirements of the marketplace are expected to increase. Shrinkage in manufacturing jobs is expected to continue, and combined with steady growth of the services sector, a persuasive argument can be made that a greater share of jobs will require strong basic skills and competencies. A recent survey of employers in New York City found that only 18 percent of jobs were open to persons without high school diplomas (Malizio and Whitney, 1984). Numerous entry-level jobs in the service sector are generated, but fewer and fewer of them will occur in industries or businesses that provide career paths for unskilled workers. Even lower-level service positions--clerical workers, cashiers, salesworkers--require basic skills in written and oral communication (National Academy of Sciences, 1984).

If high school completion rates and concomitant skill levels do not increase, "there will be a less literate pool of Americans from which colleges, universities, industry and the military will be able to draw to meet their human resource needs" (Kirsch and Jungeblut, 1986). Another prominent researcher (Levin, 1985) notes that a larger under-prepared workforce will reduce economic competitiveness, increase the cost of providing public services and produce a dual society with a large and poorly educated underclass. Kirsch and Jungeblut further predict that "this will occur unless appropriate intervention strategies are developed and implemented to meet the diverse needs of current young adults as well as to promote higher proficiencies among the younger, school-aged populations."

The factors associated with early school leaving are varied and interrelated. Some factors, such as parents' educational attainment, cannot be affected by intervention. However, two of the primary causes of dropping out--academic failure and teen parenting--are particularly amenable to intervention. The STEP program attempts to address these problems.

Academic Performance

In surveys of youth who have dropped out of school, poor academic performance is most frequently cited as the reason for leaving. Approximately 42 percent of dropouts surveyed by the High School and Beyond study reported receiving primarily Ds in their classes (Institute for Educational Leadership, 1986). A variety of research efforts undertaken in the past 15 years supports this finding. The portrait of the dropout that has emerged is one of a young person unable to function on a par with fellow students.

Beyond helping to increase the high school graduation rate, however, increasing basic skills is important because research indicates that basic skill levels in reading and math are related to employment and earnings, regardless of the individual's level of education (Berlin et al., 1986). Research on Job Corps participants and disadvantaged high school and college students has shown that well-designed remedial education programs are capable of producing modest but long-lasting academic gains (Taggart, 1981; Kapsis and Protash, 1983). In a study of black summer school students in Atlanta (Heyns, 1978), students gained, on average, the equivalent of one month for each month of instruction on the word knowledge portion of the Metropolitan Achievement Test (MAT).

Teen Parenting

The relationship between teen parenting and educational attainment is a complex one. Longitudinal studies suggest that the age at which girls first give birth is one of the strongest influences on whether they complete schooling. As many as three-quarters of female dropouts cite pregnancy and/or marriage as their reason for leaving school (Coombs and Cooley, 1968; Presser, 1978). A number of studies conclude that girls who become mothers "while they are in junior high school or high school complete on average fewer years of school, are less likely to earn a high school diploma, and are less likely to go on to college and graduate study than those who delay childbearing until their twenties" (Hayes, 1987).

The direction of causality, however, is not clear. Females who drop out of school often cite pregnancy as their reason for leaving, but there is evidence that many adolescent mothers drop out prior to pregnancy and that becoming pregnant may be precipitated by poor academic performance. As Dryfoos (1985) points out ". . . if motherhood follows unsuccessful school experiences, it may be possible that upgrading the quality of education would have an impact on fertility rates."

The consequences of adolescent parenting fall heavily on girls, but research suggests that adolescent fathers are also affected. Teen fathers are more likely to drop out of high school prior to graduation than males who do not father a child (Hayes, 1987). Their long-term education and work careers are thus likely to be affected. Data from Project Talent indicate that males who fathered a child by age 18 were over-represented in blue collar jobs compared to males who waited to until after age 18 to become fathers (Card and Wise, 1978).

A number of studies--most recently, the series of evaluations of Project Redirection (Quint and Riccio, 1985)--attest to the difficulty of successful intervention once the first pregnancy has occurred. Project Redirection researchers have concluded that, given the complexity of the problems faced by young people once they become parents, strong support should be given to program efforts aimed at preventing the first pregnancy.

Summer Learning Losses

Research conducted over the past 20 years suggests that a crucial determinant of academic performance is the extent to which learning slows, stops or even decays over the summer months. Findings from several studies indicate that this problem is particularly acute for disadvantaged youth who consistently lose ground relative to their advantaged counterparts over the summer. Studies of compensatory education programs, which serve disadvantaged youth, invariably find that participants lose ground relative to non-participants during the summer vacation (Pelavin and David, 1977; David and Pelavin, 1978; Hammond and Frechtling, 1979).

Heyns (1978) suggests that, for disadvantaged youth, actual learning regression takes place during the summer. Data from the Sustaining Effects Study show that gaps in average test scores between minority and non-minority students increased during the summer for each of the five grade cohorts in the study. Hayes and Grether (1969) suggest that as much as 80 percent of the difference between advantaged and disadvantaged students may be due to differential summer learning. These findings indicate that summer learning is a critical factor in education (Heyns, 1986).

POLICY CONTEXT

STEP operates in conjunction with the Summer Youth Employment and Training Program (SYETP), which is authorized and funded under the Job Training Partnership Act, Title II-B. The Job Training Partnership Act (JTPA), enacted in 1982, was designed to provide training aimed at increasing the employment rate and earnings of participants and reducing the rate of welfare dependency. This training was viewed by Congress as a long-term investment in human capital. In line with the position that JTPA represented a long-term investment, the legislation stipulated that 40 percent of JTPA Title II-A funds be spent on youth between the ages of 14 and 21.

The JTPA focus on youth arose from concerns about high urban dropout rates and minority youth unemployment rates. The societal consequences of the increasing under-education among youth prompted JTPA's youth expenditure requirements to ensure that adequate resources were invested in preparing youth for the world of work. In general, however, site expenditures for youth have not addressed the problems that Congress was concerned about. A two-year study of the implementation of Title II-A of JTPA (Walker, 1985) found that the training services offered were typically short-term, consisting of on-the-job training, classroom training or direct placement activities. Few services were designed especially for youth; very few jurisdictions served the youngest eligible youth (14- and 15-year-olds); and very little money was spent on remedial education. The emphasis in most sites was on meeting placement rates and unit-cost standards established by the Department of Labor. In general, the mandate to serve groups "most in need" was a low priority and sites did not expend the required 40 percent on youth programs.

Summer Youth Employment and Training Program

While Title II-A of JTPA provides funds for year-round programs for both youth and adults, Title II-B money is directed specifically toward youth for summer programs. The Summer Youth Employment and Training Program (SYETP), funded by this money, serves large numbers of disadvantaged youth, primarily through summer work experience programs.

In providing disadvantaged youth with the opportunity to earn income through a summer job, SYETP has several purposes--to "enhance the basic educational skills of youth; encourage school completion, or enrollment in supplementary or alternative school programs; and provide eligible youth with exposure to the world of work" (JTPA, Part B, Sec. 251). Although SYETP has served many youth over the years, it has realized only limited success in fulfilling these purposes.

As one of the largest national programs for disadvantaged youth, SYETP comes into contact with significant numbers of potential dropouts. SYETP services, however, have generally been limited to work experience. Although research indicates that work experience obtained during high school does increase post-high school employment, work experience is more beneficial for youth who graduate from high school than for those who drop out. Several studies have found that the most effective educational interventions for at-risk youth are those that are tied to opportunities to earn income and acquire job-related skills. One way to improve the chance of high school completion is to provide remedial instruction in addition to work experience.

Following an assessment of employment and training programs for in-school youth, the National Research Council's Committee on Youth Employment Programs recently recommended that the Summer Youth Employment and Training Program be modified to address basic skills deficiencies among the disadvantaged youth it serves. Specifically, the panel recommended that

> attempts should be made to restructure some elements of the Summer Youth Employment Program to systematically test whether SYEP [sic] elements can be used to enhance basic education sufficiently to reduce school dropout rates . . . Elements of SYEP [sic] could be structured so some skills training is added to the pure work experience in order to determine whether such training enhances the long-term employment effects of the program (Betsey et al., 1985).

Over the past few years, several changes related to the structure of SYETP have occurred. JTPA amendments passed in 1986 mandate that some Title II-B funds be spent on increasing literacy and reducing high school dropout rates. Although specific funding levels are not stipulated, the amendments require that SYETP include a remedial education component beginning in 1987. Service delivery areas (SDAs) are required to assess the reading and math skills of SYETP participants to determine whether they need remedial education. In addition, the written goals that are used to evaluate program effectiveness may include improvement in educational and employability skills.

Independent of national legislative efforts to restructure SYETP, a national study of SDAs conducted by the National Job Training Partnership in the summer of 1986 indicates that the number of SDAs offering an educational component has greatly increased over the past few years. Of the 292 SDAs responding to the survey, only 20.4 percent offered educational services in 1984. By 1986, 56.2 percent of the responding SDAs offered education and the number of youth who received educational services increased from 10,000 to 48,000. The survey results indicate that 75 percent of

SDAs expected to have an education component in operation in 1987 (NJTP, 1986).

The Summer Training and Education Program

The STEP demonstration was designed to address questions about the feasibility and effectiveness of modifying the current Summer Youth Employment and Training Program to include components that focus on some of the underlying causes of school dropout behavior. The STEP model adds two components to the summer employment program--a remediation component and a life skills and opportunities component. The remediation component focuses on the improvement of basic reading and math skills for those youth who have deficiencies in this area. STEP also addresses the interrelated problem of teenage parenting through the life skills and opportunities component. To the extent that STEP is successful in improving academic skills and decreasing the incidence of teen parenting, it addresses the issues underlying the JTPA legislation. A modified SYETP program that reduces the high dropout rates among disadvantaged and minority youth would provide the long-term investment in human capital envisioned by Congress.

ORGANIZATION OF THIS REPORT

The previous discussion of STEP's background and the policy context of the demonstration provides a framework for the remainder of this report, which covers the second summer of the demonstration in detail. The STEP model, site descriptions and a summary of early demonstration findings are presented in Chapter II. The research design used in the evaluation of STEP is described in Chapter III. Chapter IV presents an analysis of the implementation of STEP during the summer of 1986. The second cohort of STEP youth--both treatment and controls--and treatment youth in Cohort I who returned for their second summer are profiled in Chapter V. Impact analyses are discussed in Chapters VI and VII: Chapter VI focuses on impacts that can be detected over the course of one summer for Cohort II and the aggregate sample of the two cohorts; Chapter VII presents the estimated impact on school-year behavior and other outcomes that can be detected over the second summer of participation. An analysis of the costs to operate STEP is presented in Chapter VIII. Chapter IX summarizes the second year of the demonstration and discusses plans for Year Three.

II. THE STEP MODEL AND EARLY FINDINGS

The primary objective of STEP is to help at-risk youth remain in school and receive a high school diploma. STEP attempts to achieve this goal by focusing on two primary causes of early school leaving--poor academic performance and teen parenting. Remediation provides youth with an opportunity to improve their academic performance and thus their experience in their regular school. The life skills and opportunities component offers youth instruction on responsible social and sexual behavior. At the same time, youth are able to earn income that is provided by a summer job and gain work experience that may be valuable for future employment. The combination of the components in STEP is expected to increase students' motivation to complete their education.

PROGRAM COMPONENTS

With some variation in implementation across the five demonstration sites, each program focuses on four key components: remediation, life skills and opportunities, work experience and school-year support. Implementation of these components is described in detail in Chapter IV; they are reviewed briefly here.

During each of two summers, participants are offered approximately 90 hours of remedial instruction in basic reading and math skills, 18 hours of life skills and opportunities instruction, and at least 80 hours of work experience. During the intervening school year, activities designed to encourage youth to remain in school are offered. These activities are aimed at maintaining the progress STEP participants made in the first summer and encouraging them to return for the second summer.

The remediation component provides a minimum of 90 hours of skill-based group and individually paced instruction. The remediation plan requires that students spend roughly 20 percent of their instructional time in computer-assisted instruction (CAI), which is intended to motivate youth without regular access to computers. In addition, at least 20 minutes daily is spent in silent sustained reading. The curriculum was designed to address two major interlocking problems--deficiencies in basic skills and a lack of self-confidence and sense of control over the learning process.

STEP's approach to remediation--Practical Academics--is built around a set of learning "modules" that offer important advantages for youth who are performing poorly in school. Each module addresses specific reading and math skills such as categorizing and sequencing written information, distinguishing fact from opinion, interpreting charts and graphs and solving math word problems using simple and more complex calculations. The skills are taught through modeled instruction and applied, individualized follow-up activities. Exercises that require students to use selected skills are embedded in topical modules that are relevant to this population of students. Each skill is reinforced at increasing levels of difficulty; this allows students to work up to a level that challenges them and to receive prompt feedback on skill attainment and progress. Additional skill reinforcement, CAI and silent sustained reading constitute the remainder of the remediation schedule.

P/PV staff and a team of curriculum experts developed a set of remediation modules[1] that address the needs of the target population and produce essentially uniform programs at all sites. Although all modules are complete, they are designed to allow for some modification or enhancement to meet local needs. In addition to the nine modules developed by P/PV, several sites developed modules following P/PV guidelines and approval for local use.

The Life Skills and Opportunities (LSO) component of STEP stresses responsible social and sexual attitudes and behavior. Personal and sexual decision-making, job equality issues, the consequences of teen pregnancy, substance abuse, ways to avoid pregnancy including abstinence and sources of family planning services are covered in the two volumes of this curriculum. A constant focus is the relationship between decisions about sexual behavior and personal development. For example, the economic, health and social implications of early childbearing are fully explored. In addition to enhancing students' understanding of the consequences of parenting, the LSO curriculum discusses the delay of sexual involvement and/or the consistent and effective use of contraception.

P/PV, in conjunction with outside consultants, designed the curriculum to meet the special needs and circumstances of our target population, many of whom are minority group members. During the first summer, students participate in approximately 18 hours of LSO instruction. Lectures, discussions, films, roleplays, field trips and outside speakers are used to stress future planning, goals, responsibility and both the decision to have sex and how to have sex responsibly. Volume II of the curriculum, used in the second summer of the program, provides an additional 18 hours of instruction that reinforces the topics covered in the first summer and introduces new material on substance abuse, general health issues and career planning.

---

[1] P/PV developed nine thematic modules:

- two job skills modules -- Adjusting to the Working World, and Getting and Handling the Job;
- three school-based modules -- Pioneers in their Field, Early Explorers, and World Cultures--Pros and Cons of Arranged Marriages; and
- four life skills modules -- Substance Abuse, Getting the Message from the Media, Decisions, and Self-Identity.

In the work experience component of STEP, youth are placed in jobs that match, to the extent possible, their interests and skill levels. Youth are generally assigned to jobs either in their neighborhood or near their remediation site to minimize transportation problems.

STEP youth are younger than the majority of SYETP participants. Because of their inexperience and child labor law requirements, the types of jobs that are available to them are limited. Most are placed in maintenance, recreation, clerical and child care aide positions. A few youth receive assignments to special group projects such as making puppets and performing stories at local libraries and community centers. All are paid minimum wage for their participation in both work experience and remediation/life skills instruction.

The school-year support (SYS) component bridges the period between participation in the first and second summers of STEP. It serves to motivate youth to remain in school by providing counseling and referral to needed services. It also encourages youth to return for a second summer in STEP. School-year support includes both individual contacts between youth and a designated mentor/counselor and group activities held periodically throughout the school year.

School-year support has been intensified beginning with the 1986-87 school year. Each site now has a school-year support coordinator who is responsible for overall planning and implementation of SYS activities. Coordinators are responsible for hiring, training and supervising mentor/counselors and planning group activities. Mentor/counselors are responsible for making individual contacts with youth, linking them to social services as needed, and monitoring school attendance and performance.

Parental involvement in school-year support is encouraged through periodic informational mailings and meetings held at the beginning of the school year. Parents are also encouraged to assist with organizing and conducting group activities. All parents formally approve the participation of their child in STEP.

TARGET POPULATION

STEP targets disadvantaged youth who are "at risk" of dropping out of school and/or becoming adolescent parents. To be eligible for participation, youth must be 14 or 15 years old, low-income and educationally deficient but still attending school. Since STEP operates in conjunction with SYETP, youth must be at least 14 years old and economically disadvantaged to meet the SYETP eligibility requirements.

STEP is directed toward 14- and 15-year-olds, rather than older youth, for several reasons. First, students are thought to be more receptive to remedial efforts at the ages of 14 and 15 than at later ages. Second, many are making the difficult transition from junior to senior high school where problems often develop. The dropout rate begins to accelerate as youth both enter high school and approach age 16. Third, as young adolescents, they are also becoming sexually active, increasing their risk of becoming parents. Thus, their need for information regarding sexual attitudes and behavior is heightened at this point. Finally, although these youth are interested in obtaining summer jobs, their opportunities to work are limited.

Since poor academic performance is a primary reason for leaving school without graduating, the final eligibility requirement is based on academic performance. Eligible youth must perform below their grade level in reading or math, as indicated by recent standardized test scores and/or by a recent history of grade retention. However, a fourth grade reading floor is imposed to ensure that youth have the minimal decoding skills needed to benefit from the remediation curriculum.

THE PILOT EXPERIENCE

STEP was piloted in Boston, Massachusetts; Pinellas County, Florida; and Baltimore, Maryland in the summer of 1984. A total of 434 youth participated in the STEP treatment across the three sites. In Boston and Pinellas, data on control youth were also collected. A total of 322 youth were randomly assigned to the control groups. All 73 participants in Baltimore were treatment youth.

The low-income, educationally disadvantaged youth who participated in the program were primarily from minority, single-parent (usually female-headed) households. Treatment youth increased their basic reading and math skills substantially beyond what they would have in the absence of the program. Participants also gained knowledge about contraception and the consequences of adolescent parenthood. The strong results of the pilot encouraged P/PV to proceed with the full demonstration.

SITE SELECTION AND DESCRIPTION

With the decision to proceed with a national demonstration, P/PV solicited letters of interest from sites in January 1985. After an initial screening to determine each potential site's understanding of the STEP model and willingness to make the necessary commitments for program implementation, site visits were made to prospective sites to discuss their plans to operate a STEP program. Final site selection was based primarily on management and administrative experience and capacity; the cooperation of relevant institutions, most notably the JTPA agency and school district; the extent of local resource commitment; the ability to serve adequate numbers of youth, and a racial/ethnic mix that would make the results useful to other jurisdictions; and a willingness to collect the data and carry out other requirements of the research design.

As a result of the review and screening process, P/PV selected five sites--Boston, Massachusetts; Fresno, California; Portland, Oregon; San Diego, California; and Seattle, Washington. Local program operations began in the spring of 1985. On-site coordinators were hired by P/PV to serve as facilitators, program monitors, data collectors, liaison between P/PV and site operators, and providers of technical assistance to the local agencies who operate the program.

Sites soon began recruitment and enrollment of the first cohort of participants. In late June 1985, remediation, life skills instruction and work experience for Cohort I youth commenced. A full discussion of site selection, program implementation and impact analysis for the first summer was presented in a previous report (see Branch et al., 1986).

Brief descriptions of the five sites' management structure and a summary of program results for the first summer of implementation follow.

Boston

In 1986, oversight for STEP in Boston was shared by two organizations: the Mayor's Office of Jobs and Community Services (formerly the Neighborhood Development and Employment Agency) and the Boston Private Industry Council. They contracted with Action for Boston Community Development (ABCD) to operate both the Summer Youth Employment and Training Program and STEP.

Remediation was managed by three local universities: Boston University, University of Massachusetts/Boston, and Northeastern University. Treatment youth were assigned to the university closest to their home. They attended remediation classes Monday through Friday mornings, life skills classes for 90 minutes two afternoons per week and worked on campus three afternoons per week. LSO instruction was provided by ABCD counselors who also monitored work experience. Control youth worked 25 hours per week at regular SYETP jobs, most of which were in their neighborhood.

Fresno

The Fresno Private Industry Council (PIC), as the lead agency, was responsible for operational and fiscal management of STEP in Fresno. The PIC contracted with the Economic Opportunity Commission for outreach, recruitment, intake, economic eligibility determination and work site development for STEP youth. The remediation and life skills components were the responsibility of the Fresno Unified School District.

Remediation and life skills classes were held at one middle school five mornings a week. Treatment youth worked between two and a half and four hours each afternoon. Both treatment and control youth, who worked between seven and a half and eight hours per day, were assigned to jobs throughout the city at schools, parks, government offices and private businesses.

Portland

In Portland, the Portland Private Industry Council (PPIC) was responsible for all aspects of STEP. Staff hired by the PPIC to operate STEP were responsible for recruitment, eligibility determination, intake, work site development, and monitoring and managing the various components of STEP. Although teachers and aides were hired by Portland Public Schools and a school building administrator had oversight of remediation and life skills activities, day-to-day supervision was the responsibility of the PIC's STEP coordinator.

Students attended remediation and life skills classes each morning at one high school. Treatment youth worked approximately three hours each afternoon, while control youth worked six to seven hours a day. Approximately half of the jobs were custodial positions in the Portland Public Schools, but youth also worked at universities, hospitals, parks and government offices.

San Diego

The San Diego Regional Employment and Training Consortium was the lead agency with primary responsibility for STEP in San Diego. The remediation and life skills components were subcontracted to the San Diego City Schools. The work experience component was subcontracted to the San Diego Regional Youth Employment Program.

Remediation and life skills classes were held at five high schools located throughout central and southeast San Diego. Classes were held each morning, and treatment youth worked three hours each afternoon. Control youth generally worked four hours each afternoon, four days per week. With over 100 separate work sites, most treatment and control youth were placed in jobs near their home in primarily custodial, recreation, child care and clerical positions.

Seattle

Implementation of STEP in Seattle was the responsibility of the Summer Youth Employment Program in the city's Department of Human Resources (DHR). DHR had responsibility for all aspects of STEP with financial support from the Seattle-King County Private Industry Council, Seattle Public Schools, the Washington State Commission for Vocational Education, the Washington State Employment Security Department and United Way of King County.

A centrally located high school served as the site for remediation and life skills classes, which were held both mornings and afternoons. Treatment youth attended either morning or afternoon classes and worked the other half-day, while control youth worked seven hours each day. Jobs included both individual assignments in clerical, maintenance and child-care positions, and group projects in grounds maintenance, neighborhood clean-ups and theatrical programs.

THE 1985 SUMMER EXPERIENCE

The positive findings from the pilot led P/PV to expand to a three-year national demonstration of STEP starting in the summer of 1985. The five participating sites each recruited approximately 330 youth who were randomly assigned to either a treatment or a control group.

The implementation analysis further confirmed the operational feasibility of STEP. The first cohort of STEP participants was a racially diverse group who were multiply disadvantaged and clearly in need of some type of intervention. STEP was successful in attaining a number of short-term benefits for its participants in both educational and fertility-related outcomes.

The first summer of STEP was successful in stemming substantial learning losses that occurred over the summer in the absence of the treatment. Treatment youth outscored their control group counterparts in both reading and math by approximately one-quarter of a grade equivalent. The program was not powerful enough, however, to completely mitigate summer learning losses for treatment group members. Test scores at the end of the summer in both reading and math were lower than they were at the beginning of the summer. The control group losses, however, were significantly larger than those experienced by treatment youth.

Participation in STEP also had a substantial impact upon participants' knowledge of contraceptive information. Knowledge of birth control methods and availability significantly increased for every site, racial/ethnic and sex subgroup. The results also suggest that STEP was beginning to have an impact on sexual behavior. At the end of the summer, treatment group boys were significantly more likely than control group boys to say that they had not had sex in the past two months.

Positive effects, presented in this report, continued to occur during the second year of the demonstration. Before discussing the impacts of STEP in Year Two, however, we summarize the demonstration's research design in Chapter III; describe the implementation of the program in Year Two in Chapter IV; and describe the STEP participants in Chapter V. The impacts of participation in one summer of STEP on Cohort II are presented in Chapter VI, while the impacts of STEP on the first cohort of participants during the school year and their second summer of participation are presented in Chapter VII.

III. RESEARCH DESIGN

The evaluation of STEP is designed to assess both the impact of the program on participants and the feasibility of implementing the model in various settings and on a large scale.

Research began with the pilot program in 1984 and will continue until 1993, five years after completion of the four-year operational phase of the demonstration. The major components of the research are an implementation analysis and an impact analysis. The implementation analysis assesses the viability and replicability of STEP by describing the processes involved in planning, coordinating and operating the program. The impact analysis examines the impact of STEP on academic performance, development of responsible social and sexual attitudes and behavior, reduction of adolescent parenthood, school retention and high school graduation, and early labor force experience.

IMPLEMENTATION ANALYSIS

The implementation analysis examines factors that contribute to understanding the demonstration's achievements and replicating the STEP model. Topics related to initial implementation of the demonstration, such as developing and coordinating institutional arrangements, were reported previously (see Branch et al., 1986). The implementation analysis presented in this report focuses on developments related to the summer of 1986, such as:

- The precise nature of the treatment, particularly changes in the design of remediation, life skills and work experience resulting from the first summer's experience;

- Problems encountered and solutions attempted with respect to re-enrolling Cohort I, recruiting Cohort II and operating the program with a double cohort;

- Characteristics of program participants--both the new Cohort II participants and returning Cohort I treatment youth;

- Responses to key features of the program by both participants and instructors (remediation and life skills); and

- An analysis of program costs--incremental costs of operating STEP above and beyond the cost of offering summer work experience to control youth.

The implementation analysis relies upon both qualitative and quantitative data. Qualitative data are collected primarily by P/PV's local site coordinators who live in the participating communities. Site coordinators were involved in program activities associated with the initial development and implementation of STEP, and with the changes made prior to the second summer.

Throughout the summer, site coordinators spent hours observing remediation and life skills classes. They also visited and observed a sample of both treatment and control worksites. Site coordinators are well-positioned to pinpoint aspects of the program components that appear particularly effective, as well as those that are problematic. The assessments made by site coordinators are supplemented by periodic site visits and observations by P/PV's central research and program staff, and by expert life skills and remediation consultants.

The observational data are supplemented by quantitative data obtained from application forms, participants' answers to questionnaire items designed to ascertain responses to key program elements, program time sheets, payroll records and STEP budget and expenditure information.

IMPACT ANALYSIS

The impact analysis assesses the effect of participating in the STEP treatment on educational, fertility-related and work-related outcomes. Impacts are determined through comparisons of outcomes for treatment youth with those for the randomly assigned control youth who only participated in work experience through the Summer Youth Employment and Training Program. The ultimate focus of STEP is on high school graduation rates and teen parenting, but outcomes are being observed and analyzed at multiple points in time.

After one summer of STEP, we examine whether participants:

- Retain more of their reading and math skills than did the control youth, or, perhaps, improve their basic skills;

- Exhibit higher expectations with respect to educational attainment; and

- Exhibit a greater knowledge of effective contraceptive practices and the consequences of adolescent pregnancy.

Beginning with the school year following the first summer of STEP participation, we examine whether treatment youth:

- Exhibit improved attendance and academic performance during the school year;
- Exhibit reduced levels of sexual activity and/or improved contraceptive use; and
- Experience a reduction in year-to-year school dropout rates.

Over the long term, we consider whether STEP treatment youth:

- Experience a reduction in adolescent pregnancy and parenthood;
- Experience an increase in high school graduation rates; and
- Achieve increased employment and earnings in their early labor force experience.

Impact analyses presented in this report are based on the data available to us by the end of the summer of 1986. Thus, we examine the effects of one summer of program participation on youth in Cohort II (those who entered the program in 1986) and the combined (aggregate) sample of Cohorts I and II. In addition, the impacts on outcomes that occur during the school year and the outcomes after the second summer of treatment are assessed for Cohort I youth.

DESIGN STRATEGY

To answer the questions laid out above, STEP employs an experimental design involving random assignment of eligible youth to treatment and control groups. Following the determination of program eligibility and obtaining parental consent,[1] half the eligible youth at each site are randomly assigned to the treatment group, dividing their time between classes and work experience. The remaining youth are assigned to the control group, spending all of their program hours in work experience provided by the federal summer jobs program. Both treatment and control youth are compensated at minimum wage for all their program hours--both work experience and instructional time.

Random assignment is critical to the research design and the determination of program impacts. Because eligible youth are randomly divided between a treatment and a control group, the two groups are statistically equivalent in basic demographic and achievement characteristics at baseline. The only difference between the two groups is that treatment group members can participate in STEP. The experience of the control group is indicative of the experience that treatment group members would have had in the absence of the program. Therefore, we can be fairly confident that differences in outcomes that arise between the groups are attributable to the program and not to other factors.

---

[1] Written parental consent was obtained for youth to participate in a research study involving random assignment, for the release of school records data to participate in the remediation and life skills components if assigned to the treatment group.

The design calls for a combined sample of 3,000 treatment and control youth to test STEP's major impact hypotheses. This sample size enables us to be reasonably certain that we will detect policy-relevant impacts, given reasonable assumptions regarding the explanatory power of the variables in the multivariate analysis and sample attrition over the course of the follow-up period. The sample of 3,000 was built up over a two-year period with 300 youth enrolling each year (1985 and 1986) in each of the five sites. Based on the success of the second summer program operation, however, the demonstration was expanded to serve a third cohort of youth entering in the summer of 1987. The larger sample will enable us to more precisely measure the impacts of the demonstration.

DATA COLLECTION STRATEGY

Quantitative data on the individuals in the STEP sample have been and will continue to be assembled from a number of different sources and at various points throughout the course of the demonstration (see Table III.1). Among these sources are participants' performance on standardized tests, responses to self-administered questionnaires, program records, school records and questionnaires completed by remediation and life skills instructors.

Program Records

Program records, primarily application forms and timesheets, provide important information about participants and their program experiences. Application forms verify demographic information and economic eligibility for the federal summer jobs program. Program timesheets provide a record of hours of remediation and work experience. Program records are also a major source of information about the costs of operating STEP.

Academic Testing Over the Summer

The reading and math subtests of the Intermediate Level Metropolitan Achievement Test (MAT) Survey Battery provide the primary measures of academic performance over the summer. These subtests, consisting of 60 items in reading comprehension and 50

Table III.1

SUMMER TRAINING AND EDUCATION PROGRAM DATA BASE BY SOURCE

Variables	Data Source
Demographic Characteristics Age Sex Race/Ethnicity Household Composition AFDC Receipt Single-Parent Household Marital Status Public Assistance Receipt Spanish Surname Language Spoken at Home	Application Forms School Records Baseline Questionnaire
Academic Performance Reading Mathematics	Metropolitan Achievement Test (MAT), Survey Battery, Intermediate Level I, Forms JS and KS
Education Highest Grade Completed Ever Repeated a Grade Grade Repeated Number of Days Absent in Previous Year School of Attendance at Intake Enrolled or not Transferred from Previous Year's School School Transferred to Attendance (excused and unexcused absences) Grade Point Average Grade Retention/Progression Date of Graduation Date Officially Withdrawn Reason for Withdrawal Education Aspirations Standardized Test Scores	Application Forms School Records Baseline Questionnaire Follow-up Interviews

Table III.1 - (cont'd)

SUMMER TRAINING AND EDUCATION PROGRAM
DATA BASE BY SOURCE

Variables	Data Source
Fertility/Family Planning	Baseline Questionnaire
Percent Sexually Active	Endline Questionnaire
Age at First Intercourse	Follow-up Interviews
Contraceptive Knowledge	
Contraceptive Utilization	
Knowledge of Consequences of Adolescent Pregnancy	
Number of Pregnancies	
Age at First Pregnancy	
Outcomes of Pregnancies	
Number of Children Fathered	
Numbers of Births	
Date of Birth for Each Child	
Child Care Assistance	
Employment	Baseline Questionnaire
Previous Employment	Endline Questionnaire
Career Aspirations	Follow-up Interviews
Attitudes Toward the Work-Roles of Men and Women	
Employment Status and History	
Hours per Week	
Type of Occupation	
Earnings	
Welfare Assistance	
Unemployment Compensation	

items in math, provide nationally normed performance measures. The intermediate level test is most appropriate for students performing at the fifth- to seventh-grade level. The intermediate level was selected based upon the finding that STEP youth in the pilot study generally attained scores in this range at baseline.

Once random assignment is completed, both treatment and control youth are administered the MAT to assess baseline reading and math abilities.[2] Youth are re-tested at the end of the first summer. To protect against memory effects, participants are administered one form of the MAT as a pre-test and another form as a post-test. These two forms--Forms JS and KS--are judged to be functionally equivalent. In other words, the score attained is independent of which form is taken.

Although the initial research design called for testing both treatment and control youth in the second summer, the experience of the pilot study resulted in a revision of this plan. The pilot experience indicated that without the offer of employment during the second summer, too few control youth returned for testing to justify the expense and effort involved in locating, recruiting and testing controls on a demonstration-wide basis. Only San Diego decided to offer control youth a second summer of work experience; test score and questionnaire data were not collected at the end of the second summer for controls in other sites. However, school records and youth interview data will be collected for all treatment and control youth during the longitudinal phase of the study.

In order to minimize the possibility that control youth have less motivation to perform well at the post-test than treatment youth --and thereby spuriously elevate the estimate of program impact --an incentive scheme was devised to provide both groups with an external reason for doing their best. Any youth taking the test received one chance in a drawing for one of fifteen $50 prizes awarded at each site. S/he received an additional chance for each correct answer scored on the MAT. Thus, a youth could maximize his/her chances of winning by trying to correctly answer as many questions as possible. A youth responding at random, or drawing designs on the answer sheet, or simply not answering the questions would decrease his/her chances of winning. This incentive was available to both treatment and control youth at both the pre-test and post-test. If these incentives have their predicted effects, any control group losses observed should be the result of true losses in academic skills.

---

[2] Although random assignment occurred before testing, neither proctors nor students were informed of its outcome until after testing.

In addition to providing incentives to maximize students' motivation to perform well on the tests, care was taken to ensure that testing conditions were identical for treatment and control youth. Sites were required to use the same locations and proctors for treatment and control youth. In most cases, treatment and control youth were tested side-by-side in the same classrooms.

Questionnaires

Self-administered questionnaires, completed in conjunction with testing at the beginning and end of the first and second summers, provide data on demographic characteristics and on measures of attitudes, knowledge and behavior that are relevant to the program. Background data collected include household composition, languages spoken, grade enrollment, retention and absenteeism. Attitudes and knowledge measured include educational expectations, occupational/career aspirations, awareness of the consequences of adolescent pregnancy and birth control knowledge. Responses to various components of STEP were measured using the endline questionnaires.

Teacher Questionnaires

To supplement the data collected from youth, remediation and life skills instructors completed questionnaires at the conclusion of the 1986 summer program. Teachers were asked to provide basic demographic information on themselves, such as racial/ethnic, sex and educational background. The questionnaires measured the teachers' experience with the population served by STEP and their experience with individualized teaching methods, as well as their general response to the program. These data are merged with the individual student records so that differential impacts due to characteristics of teachers can be detected.

School Records

Data from school records are essential for determining the long-term effects of program participation on school performance, retention and high school graduation. School districts in the participating demonstration sites agreed to provide school records data including standardized test scores, attendance, grade progression and dates of withdrawal, transfer and/or graduation for each youth in the STEP sample for the duration of the research phase of the project. In addition, the schools have agreed to provide data on local policies and procedures to help interpret student data and to lend assistance in tracing students and gaining access to their records when they transfer to other school districts.

The collection of school data began at the end of the 1985-86 school year for all Cohort I treatment and control youth. The school districts provided information on each youth who was still in its district. For students who transferred out of the school district, before or during the 1985-86 school year, the old school sent us the name and address of the student's new school. P/PV then contacted the new school to obtain the information required. If the original school did not know where the youth was, P/PV tried to locate the youth by calling contact people, such as parents or other relatives. Through these tracking methods, we were able to find many of the youth and identify their school enrollment status in order to collect the necessary data. P/PV will continue to track all students in the STEP sample and collect school data through 1993.

Follow-up Interviews

After the operational phase of the demonstration, STEP youth from both the treatment and control groups will be followed-up three times through telephone interviews. Data collected through these interviews will address issues of the long-term impact of STEP participation that cannot be determined from school records data. The interviews will gather retrospective information on sexual activity and contraceptive use, pregnancies, children fathered, births, employment and earnings history, current employment status and earnings, and verification of educational data including current status, attendance, grade progression, educational aspirations and high school graduation.

Summary

In summary, data are being collected both during the operational phase of the program and the longitudinal follow-up study of STEP youth. During the operational phase, the primary data sources are the summer tests, questionnaires and program records. All randomized youth are scheduled to attend testing sessions at the beginning and end of the first summer of STEP. Treatment youth who return for the second summer attend additional testing sessions at the beginning and end of the second summer.

Longitudinal data collection began with the collection of school records data on the first cohort for the 1985-86 school year. School data for all treatment and control youth will be collected annually through 1993. Personal interviews will be conducted with all treatment and control youth every two years, beginning in the fall of 1987.

The data that were available to analyze for this report include the baseline and end of first summer test and questionnaire data for both Cohort I and Cohort II, beginning and end of second summer test and questionnaire data for returning Cohort I treatment youth (and control youth in San Diego), and 1985-86 school-year data for Cohort I.

ANALYTIC STRATEGIES

Estimation of the impact of participation in STEP relies heavily on multivariate analysis. In most of the analyses, regression analysis provides estimates of the differences between treatment and control youth on outcome measures controlling for differences in important baseline and demographic characteristics. Given the randomized design, the differences in mean outcome measures provide unbiased estimates of the program impacts. However, multiple regression analysis produces more precise estimates of the effects by incorporating more information into the estimation process. Although the treatment and control groups are very similar, differences can arise either by chance or because of different patterns of attrition. Regression analysis controls for such differences on baseline characteristics. In addition, to the extent that outcomes are related to initial characteristics, regression models explain some of the variation between individuals. By controlling for the variation due to these baseline variables, it is possible to more precisely attribute the remaining unexplained differences in means to the program.

A similar specification is applied to most of the regression models used in analyzing the impact of participation in STEP. The explanatory variables include: basic demographic characteristics such as sex and race/ethnicity; demonstration site; variables that capture the youths' home environment, e.g., family size and birth order; and educational characteristics such as absenteeism and educational expectations.

Learning and school data analyses models contain additional education-related characteristics including grade or last grade attended, whether the youth has ever been retained a grade, enjoyment of reading and math, and baseline Metropolitan Achievement Test (MAT) scaled scores in reading and math.[3] Data that describe characteristics of the STEP teachers were also collected for the 1986 program and were included in the learning analysis for the second cohort.

Models used to estimate the impact of STEP on contraceptive knowledge and use, consequences of adolescent pregnancy and delay of sexual initiation include baseline measures of these same fertility-related variables, in addition to background and demographic variables.

In general, the multivariate models used to estimate the impact of STEP on various outcome measures take the following form:

---

[3] The baseline reading score is included in both the reading and math equation while the baseline math score is included only in the math equation.

(1) $Y_2 = a + b_1 Y_1 + b_2 X + b_3 T + e_1$

where: Y_2 = the post-program (endline) value of the variable of interest

Y_1 = the pre-program (baseline) value of the variable of interest

X = a vector of explanatory variables

T = whether or not the youth received STEP treatment

a, b_i = coefficients

e_1 = a stochastic disturbance term with a mean of zero and a constant variance

This specification makes it possible to control for pre-existing differences among youth while the impact of STEP is estimated (b_3).[4]

The overall impact of STEP is estimated by the dichotomous variable, T, indicating control group status. In addition to estimating the overall effect of the program using equation (1), a series of subgroup-treatment interaction variables are used to estimate the effect of the program in particular sites and to estimate the effect of STEP on racial/ethnic and sex subgroups. Algebraically, equation (1) is modified as follows:

(2) $Y_2 = a + b_1 Y_1 + b_2 X + b_3 T + c_1 TS_1 + \ldots + c_4 TS_4 + e_2$

(3) $Y_2 = a + b_1 Y_1 + b_2 X + b_3 T + c_1 TR_1 + c_2 TR_2 + c_3 TR_3 + c_4 TM + e_3$

where: S_i = site dummy variables[5]

---

[4] This model is a more robust specification than one that analyzes change scores. An analysis of change scores assumes that the amount of change and pre-test level of the outcome measure are perfectly related. If that assumption is violated, an analysis of change scores is a mis-specification of the model and the resulting estimates of the coefficients are incorrect. The model estimated for the analysis reported here controls for pre-test level if this assumption is violated and is equivalent to the change score model if this assumption holds.

[5] One site--Seattle--is omitted from the equation.

R_i = racial/ethnic dummy variables[6]

M = a dummy variable that equals 1 for males

Using equations (2) or (3) the estimated impact on any individual can be calculated. A subgroup impact (such as that for Hispanics) presented in the report is the average of each subgroup member's individual impact, assuming subgroup members differ from the rest of the sample only by being a member of that subgroup (i.e., their other characteristics, such as the percent male, are the same as the average in the sample).[7]

To further explore the effects of STEP on learning outcomes, the models of post-test scores are estimated replacing the dichotomous treatment variable, T, with the number of classroom hours a student attended. This analysis provides an estimate of the average hourly impact of STEP on learning outcomes. Estimating this model using ordinary least squares (OLS) is inappropriate and would probably produce a biased estimate of the program's impact. Unmeasurable characteristics that influence how well a participant does on the post-test (such as motivation) may also influence how well the youth attended the program. Instead of OLS a two-stage least squares technique is used to estimate models that include hours.

The use of OLS is also not warranted when the dependent variable is a dichotomous variable, such as whether a participant uses contraception during the program period or becomes a school dropout.[8] In such cases, a nonlinear maximum likelihood estimation technique, logit, is used to estimate treatment effects.

The key hypothesis of interest for our analysis is whether STEP has an effect on various outcome measures. In our discussion of the results, we indicate whether an impact estimate is statistically different from zero by labeling statistically non-zero estimates as "significant." In this text, this term is reserved for estimates that are not equal to zero at 0.10 or greater level of significance.

---

[6] One racial/ethnic group dummy--whites and others--is omitted.

[7] This estimate is calculated as an appropriately valued linear combination of treatment and treatment-interaction coefficients. In particular, the estimated impact on subgroup R_1 is: $b_3 + c_1 x 1 + c_2 x 0 + c_3 x 0 + c_4 x$ (the average percentage of males among all treatments).

[8] See Amemiya (1981) for details about the problems involved in estimation with dichotomous variables.

When discussing subgroup estimates, a second hypothesis is of interest: does the effect of STEP differ with respect to a particular characteristic, such as race or ethnicity. A test of whether the subgroup impacts differ from one another was conducted for all the subgroup analyses.

In summary, a variety of analytic strategies are used to evaluate the impact of participation in STEP. The fundamental approach used a dummy variable (indicating treatment or control group status) in ordinary least squares regression. Other analyses are used where the assumptions of the ordinary least squares model are likely to be violated; where dependent measures are jointly determined; where the measurement of the treatment variable needs refining; and where program impacts for selected subgroups of participants are examined.

IV. IMPLEMENTING STEP IN THE SECOND SUMMER

STEP implementation requires close collaboration among a variety of institutions--private industry councils, school districts and employment and training agencies--at a level of intensity and complexity that is new to most sites. The public school systems, in particular, are involved to a degree not generally characteristic of summer youth employment programs. Involvement of school districts ranges from taking charge of the entire remediation component, to providing work slots, to offering space and equipment. The school systems also "certify" the instructional component of STEP by granting academic credit for successful completion. Overall, collaboration in the second year of the demonstration was stronger than it was in Year One.

The second summer of the STEP demonstration showed improved planning and increased operational efficiency in virtually all areas, including participant identification and selection as well as the actual implementation of the three major program components. This chapter describes the activities associated with enrolling Cohort II and re-enrolling Cohort I youth; operating the remediation, life skills and work experience components in the summer of 1986; and providing school-year support to Cohort I youth during the 1985-86 school year. (Readers are referred to Branch et al., 1986, for a full discussion of STEP implementation in the first summer.)

PARTICIPANT SELECTION

This section describes how the target population for STEP was defined, how participants were recruited and how the random assignment procedures were conducted.

Eligibility Criteria

The standard eligibility criteria established for STEP at its inception remained intact during the second summer. To be eligible, youth had to be 14 or 15 years old on the first day of the program, JTPA-eligible and judged academically at risk of dropping out of school as measured by performance on standardized tests. To insure that all youth accepted into the program would be capable of understanding the information in the STEP curriculum, a minimum fourth grade reading level was stipulated as an initial screening measure in all sites. Students whose school district standardized test scores were below the fourth grade floor were, therefore, not eligible for selection.

Sites defined the general STEP criteria in terms of their local standardized tests. These criteria are shown in Table IV.1. Boston, Portland and San Diego maintained the same definitions in

Table IV.1

ACADEMIC ELIGIBLITY MEASURES BY SITE

Site	Primary Academic Screen	Secondary Screens
Boston	MAT score between the 20th and 40th percentiles in reading or math and/or Degree of Reading Power Test score of 25 to 47	Preference initially given to 8th graders
Fresno	At least one grade level below current level in reading or language or math as determined by the California Achievement Test (CAT)	None
Portland	At least one but not more than four years below grade level in reading and/or math using the Portland Area Levels Test (PALT)	None
San Diego	One to four years behind the grade norm using the California Test of Basic Skills (CTBS)	None
Seattle	One grade level below current grade in reading or math using the California Achievement Test (CAT) or the Wide Range Achievement Test (WRAT)	Retained in the past five years

the second summer that were used in Year One. The other two sites--Fresno and Seattle--made minor modifications to their definitions.

Fresno used only one test to screen for academic eligibility: the California Achievement Test (CAT). The Basic Skills Achievement (BSA) test, which was also used in the first year, was eliminated from the screen. The decision to eliminate the BSA was made because its scoring did not adequately detect reading levels below the minimum fourth-grade floor. Seattle modified the criteria to include performance at one or more grades below the student's grade level in math as well as in reading (based on the CAT or the Wide Range Achievement Test), and the grade retention criterion was restricted to the past five years. In the first summer, Seattle selected only youth who performed one or more grade levels below in reading or who had been retained at any time in the past. Portland also established additional screens for selection as a backup, but was able to recruit sufficient youth using the original criteria.

The process of identifying and recruiting youth for a second cohort of STEP was greatly improved in the second year. Sites developed lists of youth meeting the academic criteria earlier and began the entire recruitment process earlier as well. However, despite targeted outreach, it remained difficult to recruit sufficient numbers of youth who met both the academic and economic criteria. That is, although most sites started out with lists containing more than five times the number of academically eligible youth required for STEP, difficulty in recruiting persisted because many of the youth listed never applied to SYETP for jobs. To illustrate, Boston's lists of academically and age-eligible youth contained a total of 2,466 youth; Fresno's 1,955; Portland's, 1,876; San Diego's over 1,800 youth; and Seattle's 1,706. As discussed below, all sites had to intensify their planned recruitment efforts to achieve the enrollment goal despite these numbers.

Outreach and Recruitment Strategies

In the second summer of STEP, sites were involved in two separate recruitment efforts--recruitment of a new cohort of treatments and controls (Cohort II) and re-enrollment of treatment youth from the first cohort.[1] Recruitment of youth for a program that combines both work and remediation creates barriers not present in the regular SYETP system. Among the recruitment barriers that should be expected are: the necessity of establishing educational as well as economic eligibility; serving younger youth than are traditionally served by JTPA; and the reluctance of some youth to engage in summer study. Strategies for recruiting both cohorts are discussed separately below.

---

[1] In San Diego, both treatment and control members from Cohort I were invited to participate in a second summer of STEP.

Cohort II Recruitment

As was the case in the first summer, schools and agencies administering SYETP worked together to identify youth eligible for STEP based on the age, academic and economic eligibility criteria specified by P/PV and refined locally. This process of identifying and recruiting STEP-eligible youth was one of several areas identified as requiring an earlier start. In general, these activities did start well in advance of the previous year's schedule. One site developed lists of academically eligible youth as early as December. Portland did not have an accurate list from which to work until early April, but all the other sites had outreach activities underway by early March.

As noted earlier, outreach and recruitment activities again consumed a great deal of time and other resources, but were typically carried out more efficiently than in the first summer. The recruitment effort was very successful; a total of 1,635 youth were randomized to the Cohort II sample of STEP, an increase of 42 over the previous summer. The sites used more targeted and personalized strategies to recruit youth for STEP in the second summer.

Targeting involved a more active role for school principals, guidance counselors and/or other school staff. Since most lists categorized students by school, school staff were asked to encourage listed youth to apply to the program. Other enhancements to the recruitment effort included earlier meetings at schools with potentially eligible youth and school staff, earlier distribution of application packets, more direct mailings to parents and youth, and calls and visits to the homes of youth.

As a result of the earlier starts and more targeted recruitment approaches, recruitment of STEP youth was completed earlier during the second summer in Boston, Fresno and Seattle. In San Diego, the majority of youth were accepted into the program more than one week prior to the start of the program, though some were accepted up to two days following the official start. Portland had to postpone the start of the program by one week to achieve the enrollment goal.

Random Assignment

As discussed earlier, an integral part of the STEP research design is the random assignment of eligible youth to a treatment or control group. Using procedures developed by P/PV, site coordinators were responsible for carrying out the random assignment. During Year Two, additional screening to eliminate youth already randomized to Cohort I was required. Because of the sporadic pattern of application experienced in the first summer, random assignment of Cohort II youth was generally conducted in two rounds--randomizing roughly half of the sample each time. Splitting the process in this way allowed for a more timely placement of youth in appropriately scheduled jobs as well as earlier notification for the youth themselves.

Random assignment followed verification of eligibility and receipt of signed parent and youth consent forms. While youth were informed of their selection for STEP and a summer job as soon as the eligibility process was complete, the specific assignment to treatment or control status was to be withheld until the completion of the pre-test.

Re-enrollment of Cohort I

All youth randomized to the treatment group in the first summer are eligible to participate in a second summer of STEP. (In San Diego, control youth were also offered the opportunity to return for a second summer of work experience.) Youth who were no longer JTPA-eligible were also allowed to continue in the second summer. P/PV and/or local STEP agency funds covered the costs of serving JTPA-ineligible youth. Efforts to re-enroll youth began as early as February in Fresno, but in general commenced in March.

Based on the pattern of participation in the first year, Cohort I treatment youth can be categorized as follows: **school-year support participants**--those youth active in the academic year component of STEP; **program completers**--those youth who participated in at least 80 percent of the first summer's activities, but were not active in the school-year support component; **program non-completers**--those youth who participated for less than 80 percent of the first summer's activities and were not active in the school-year support program; and **no-shows**--those youth randomized to STEP who never participated in any program activity. Youth in all categories were eligible for a second summer unless some serious negative behavior barred their further participation. The strategies for re-enrolling these groups required some variation as discussed below.

In most sites, the School-Year Support (SYS) program was the primary mechanism for encouraging school-year support participants to re-enroll for a second summer. SYS staff reminded youth of the guarantee of a job in the upcoming summer and stressed improvements in the program for the second year, such as better curriculum, better jobs and, in the case of Boston, a raise of 20 cents per hour. Unfortunately, participation in school-year support was disappointing during the 1985-86 school year, thus school-year support did not have much influence on re-enrollment rates. In San Diego, which had no school-year support component, treatment youth returned for a second year of STEP in proportions similar to treatment youth in Fresno, Portland and Seattle, where school-year support was offered.

Some sites also used school-year support staff to augment JTPA agency efforts to recruit program completers, non-completers and no-shows. SYS staff sent numerous letters, made phone calls and invited Cohort I youth to re-enrollment events. In addition, non-completers and no-shows in Boston were required to attend a personal conference with an ABCD counselor to re-establish the original agreement governing their participation in STEP.

As a result of intensive and often personal contacts, almost 80 percent of the original Cohort I treatment sample re-enrolled. Boston had the lowest return rate (68 %), and Fresno had the highest rate with 85 percent of their treatment youth re-enrolling (see Table IV.2). Portland re-enrolled 83 percent of its Cohort I treatment sample, while both San Diego and Seattle re-enrolled 79 percent. In San Diego, where control youth were eligible to participate, 89 percent (or 10% more controls than treatments) of the original control sample re-enrolled for a second summer of work experience.

Common reasons given by Cohort I youth for not returning included the need to go to summer school for specific credits, better paying jobs, and the preference for a summer job over a summer job coupled with school. In Boston, the low return rate was attributed largely to the healthy economy and attendant job opportunities for youth. In addition, being one year older with some work experience is believed to have opened up more opportunities for Cohort I youth in Boston's booming economy. Nevertheless, about four out of five treatment youth did return for a second summer of STEP.

TESTING

Forms JS and KS of the Survey Battery, Intermediate Level of the Metropolitan Achievement Test (MAT), were interchanged for pre-testing and post-testing the ability of STEP youth in reading and math. Along with the MAT, a questionnaire was administered to collect data on demographic characteristics, educational history, aspirations, work history, birth control knowledge and fertility-related characteristics. Slightly different versions of the questionnaire were given to Cohort I and Cohort II youth. P/PV's site coordinators trained test proctors in administration procedures.

Based on the experience in Year One, P/PV made several improvements in the procedures to ensure that treatment and control youth were tested precisely in the same way. First, we enforced the requirement that both treatment and control groups receive equivalent testing conditions. This ensures that any existing

Table IV.2

RE-ENROLLMENT OF COHORT I YOUTH IN 1986

Site	Number of Treatment Youth Randomized in 1985	Percent Re-Enrolled in 1986	Number of Control Youth Randomized in 1985	Percent Re-Enrolled in 1986
All Sites	798	79%	796	na
Boston	165	68	165	na*
Fresno	158	85	160	na*
Portland	155	83	155	na*
San Diego	156	79	154	89%
Seattle	164	79	162	na*

*Control youth were not offered a second summer in these sites.

na = means not applicable

testing irregularities or systematic biases in the testing situation would be equally experienced by both groups. In Year One, treatment and control youth were often tested separately. However, during the second summer, they were tested in the same classrooms, with the exception of Fresno. In Fresno, local restrictions prevented strict adherence to the requirement of testing both study groups together, but the same proctors and classrooms were used for all test administrations.

The consistency of proctoring and of test locations was also improved in the second year of STEP. During the first summer, staff from the agency administering SYETP often proctored exams for control youth, while teachers were responsible for testing treatment youth. In Year Two, however, STEP teachers were responsible for proctoring testing sessions for both treatment and control youth in most cases. Again, due to local restrictions in Fresno, non-teaching staff proctored exams for both treatment and control youth. Finally, during Year One, many sites tested treatment youth in classroom settings while control youth were generally tested in other settings designed to accommodate larger numbers of youth. Except for make-up testing sessions, treatment and control youth were all tested in classrooms for the second summer of STEP.

In addition to the changes noted above, sites also made local logistical changes designed to facilitate a more consistent testing operation. P/PV hired local observers to view a sample of classrooms at each site and report on the process. Observers completed structured instruments, and site coordinators provided a narrative report designed to provide fuller insight into the local testing situation. In general, testing conditions and procedures improved over Year One. However, isolated incidents of poor testing situations point to the need for greater intervention on the part of test observers when poor conditions occur and for better training of both proctors and observers.

Efforts to test a high percentage of youth were as successful in the second summer as in the first. As Table IV.3 shows, across all sites, 90% of Cohort I youth who re-enrolled took the pre-test, as did 93% of youth randomized to Cohort II. Similarly, across all sites, 91% of the re-enrollees who were pre-tested took the post-test compared to 92% of Cohort II youth. The general patterns that were found are discussed below.

Pre-testing

In general, youth took the testing situation seriously, with a few exceptions: in Boston and Fresno, where treatment youth from both cohorts were tested separately, returnees showed a less serious attitude and were less attentive than the incoming cohort. There were isolated incidents of cheating or attempts at cheating at four of the five sites, and notification of treat-

Table IV.3

YOUTH TESTED BY SITE FOR BOTH COHORTS

	All Sites	Boston	Fresno	Portland	San Diego	Seattle
Cohort I						
Number of Treatments Randomized in 1985	798	165	158	155	156	164
Number (%) of Treatments Re-enrolled	628 (79%)	112 (68%)	134 (85%)	129 (83%)	124 (79%)	129 (79%)
Number (%) of Re-enrollees Pre-tested	564 (90%)	88 (79%)	124 (93%)	122 (95%)	109 (88%)	121 (94%)
Number (%) of Pre-tested Re-enrollees Post-tested	515 (91%)	84 (95%)	109 (88%)	114 (93%)	101 (93%)	107 (88%)
Cohort II						
Number Randomized	1635	331	328	325	321	330
Number (%) of Randomized Sample Pre-tested	1519 (93%)	311 (94%)	318 (97%)	303 (93%)	293 (92%)	294 (89%)
Number (%) of Pre-tested Sample Post-tested	1393 (92%)	284 (91%)	276 (87%)	271 (89%)	277 (95%)	285 (97%)

ment/control assignments occurred in one classroom in Portland prior to test administration. (Instructions to proctors require that group assignment be withheld until after the test is taken to guard against any differential reaction to the test by treatment and control youth).

Post-testing

Youth were less attentive at the post-test than at the pre-test. For the sample of classrooms observed across sites, observers described the attitude of youth at the post-test as inattentive in all but two classrooms in San Diego. In addition, notations of apathy and boredom were made in all sites. Again in Boston and Fresno, more inattentive behavior was noted for Cohort I youth. Treatments and controls were also tested together for the post-test, and the less attentive behavior was observed for both groups. This observation underscores the importance of examining the net difference between treatment and control scores, rather than within group change scores since the two groups appeared to behave similarly in the testing situation.

THE REMEDIATION COMPONENT

Remediation in reading and math is central to the STEP treatment. During the first summer of the full demonstration, P/PV developed a curriculum framework called CORE (Community Organized Remedial Education) for use by sites in developing locally relevant curricula. In brief, the CORE curriculum framework specified 14 competency areas--five in reading and nine in math. Within this framework, sites were to develop a curriculum that met local needs by defining local competencies and objectives in reading and math, selecting initial diagnostic locator tests, deciding on appropriate instructional strategies, choosing instructional materials and assessment exercises, preparing forms to document student progress and pulling together the logistics necessary for implementing the curriculum.

Sites were surprised at the amount of time and responsibility required to develop the curriculum that this approach necessitated. Further, CORE fell short of overall site expectations, and problems with its implementation led P/PV to rethink and redesign the approach in Year Two. Maintaining the basic instructional approach of a skill-based, individually paced curriculum, P/PV's central and consulting staff developed a scripted curriculum, called Practical Academics, for use by sites during the 1986 summer.

Curriculum Description

The curriculum, which calls for local refinement as needed, retains the key elements of the first summer's design as follows:

an approach that centers instruction on a set of defined competencies or skills; an educational program tailored to the individual learner's needs; a minimum of 90 hours of instruction; computer-assisted instruction (CAI) for 20 to 25 percent of instructional hours; a teacher-student ratio of at least 1:15; and remediation sessions not exceeding three hours daily (excluding breaks).[2]

The biggest difference between CORE and Practical Academics is that Practical Academics is an actual scripted curriculum for use by sites. The Practical Academics curriculum also requires 20 minutes of silent sustained reading as part of the daily remediation experience. A sample weekly schedule for the classroom component of STEP is presented in Table IV.4.

The curriculum includes teacher-led group instruction, student-paced skill reinforcement exercises, opportunities for supplementary projects (group and individual), and independent reading and writing assignments. This approach was selected because it combines direct instruction with focused practice using motivational material tailored to individual needs. Furthermore, while many remedial learners can master skills in drill exercises, they are often unable to transfer their new skills to school, job or life situations.

At the center of the Practical Academics curriculum are topic-based learning modules that provide basic skills instruction in the context of issues of particular relevance to high school students. P/PV developed nine such learning modules. Three focused on school-based skills: Pioneers in the Field, Early Explorers, and World Cultures--Pros and Cons of Arranged Marriages. Four focused on life skills: Substance Abuse, Getting the Message from the Media, Decisions, and Self-Identity. And two focused on job skills: Adjusting to the Working World and Getting and Handling the Job.

Each module is designed as a week-long unit and consists of three subtopics. Brief initial instruction within each module is designed for presentation to the entire class, but applied activities are developed to cover the range of expected ability levels in a STEP classroom. Activities were developed at basic, intermediate and advanced levels. These levels are different-

---

[2] Remediation was held during the morning hours in all sites, except Seattle where both morning and afternoon sessions were held for Cohort II youth. Cohort I youth in Seattle attended remediation only in the afternoon. During Year Three Seattle will offer remediation sessions only in the morning.

Table IV.4

SAMPLE WEEKLY SCHEDULE
LEARNING MODULE: GETTING THE MESSAGE FROM THE MEDIA

	MONDAY Topic: Newspapers	TUESDAY	WEDNESDAY Topic: Consumer Ads	THURSDAY	FRIDAY Topic: Using the Telephone Book
8:30 - 9:15	Whole Group Presentation 1. Motivational Activity 2. Modeling and Direct Instruction 3. Preparation for Applied Activities	Life Skills & Opportunities (Separate Curriculum)	Follows Monday's Schedule	Life Skills & Opportunities (Separate Curriculum)	Follows Monday's Schedule
9:15 - 10:00	Applied Activities 1. Independent/3 levels 2. Task-oriented activities which incorporate selected reading and math skills				
10:00 - 10:15	BREAK	BREAK	BREAK	BREAK	BREAK
10:15 - 10:30	Summary of Lesson Whole Group Presentation	Completion of previous day's activities	Follows Monday's Schedule	Follows Tuesday's After-Break Schedule	Follows Monday's Schedule
10:30 - 12:00	Additional Student Activities *1. Individualized instruction and practice of reading & math skills. and/or *2. Suggested independent projects and 3. Related CAI (equivalent of 30 minutes each day, minimum) and 4. Independent reading/writing				

*May include small group instruction.

iated by the level of readability and by the complexity of the skills required for completion.

Three modules were recommended for use in the first summer, three were recommended for use in the second summer and the remaining three modules could be used in either summer. In practice, there was little uniformity across sites in the assignment of the modules to cohorts. At each site, a cohort used approximately six modules over the course of the summer.

In addition to the modules developed by P/PV, each site also developed at least one module locally. Boston developed a module in the school-based category, called the Literature Connection, in which youth read one of two novels. San Diego developed a supplement to its school-based modules using the newspaper. Fresno and Seattle developed modules in the job skills area. In Fresno, the module was titled Communication on the Job, and Seattle developed Jobs of the Future. Boston modified the P/PV job skills module, Getting and Handling the Job, to make it more relevant to their local program and called it Your Summer Job. Portland was the only site that developed a module in the life skills area, called Family Law.

Skill reinforcement is another key element of the Practical Academics curriculum. The skill reinforcement portion of the curriculum provides students with an opportunity to practice individual skills modeled through instruction in the learning modules. Skill reinforcement provides a further opportunity to tailor instruction for individual students, as well as an opportunity for use of CAI in reading and math skills.

Implementation of the Remediation Component

The use of topic-based learning modules as the organizing principle for providing remedial instruction was well received. The prescriptive approach of Practical Academics was preferred over the "local option" approach of CORE. The overall implementation of remediation, including the provision of CAI, was better planned and ran more smoothly in 1986 than in 1985.

However, because the module approach to remediation was introduced for the first time in the summer of 1986, some start-up time was again required. Teachers needed time to become familiar with the learning content of the modules, to get comfortable with the script and discover the points at which they might want to vary it in class, and to plan how they would coordinate reinforcement activities and CAI with the new modules.

Several implementation changes strengthened the component in the second summer. These changes included providing increased consulting assistance in remediation planning by P/PV, hiring a lead teacher in each site, better screening of remediation teachers, earlier and improved training of teachers, and changes made in the use of CAI. These are discussed in detail below.

Lead Teacher

In February, each site hired a lead teacher responsible for overall management of the remediation component. This appointment was designed to give remediation full-time direction in each site and to cultivate increased and consistent involvement by the schools--areas found to be deficient during the first summer.

The lead teacher was responsible for organizing a curriculum development team of reading and math teachers to develop a local module, reviewing the P/PV-developed modules, identifying print and CAI materials, and planning the overall remediation component. The lead teacher also worked with members of the curriculum development team and P/PV to develop a training program for STEP teachers.

Overall, the designation of a lead teacher as educational liaison for STEP gave the remediation component higher priority and an increased level of attention. However, the level of authority enjoyed by lead teachers and the competency of lead teachers varied by site. For example, the level of effort in curriculum development and refinement varied both in scope and quality across sites as did the ability of lead teachers to obtain equipment and material resources, and provide teacher training and supervision.

Consulting Assistance

In Year Two, increased technical assistance was provided to the sites. Two consultants and P/PV's in-house educational staff shared responsibility for providing assistance on remediation planning and management. Five days of consultant time were provided to each site between February and the end of the summer. In general, consultants assisted in developing the local module, refining the P/PV-developed modules, developing the instructional plan and training teachers by meeting with key site staff.

The effectiveness as well as the amount of technical assistance also increased, since the lead teacher served as a consistent contact with consultants. The consultants, in turn, were more effective when their expertise filled a specific need and when they were involved in a specific project from beginning to end.

Teacher Selection and Characteristics

To improve the average quality of the remediation teachers, P/PV developed guidelines for sites to use in selecting teachers. Qualifications included backgrounds in either reading or math and experience with disadvantaged populations. In addition, an ethnically and racially diverse group of teachers was sought so that there would be a greater congruence between the racial and ethnic backgrounds of the teachers and students. The process of screening STEP teachers began earlier in the 1986 summer than in the first year. The overall implementation schedule called for all teachers to be hired by May 1st. Below, we examine how well the characteristics of the teachers fit the guidelines.

Sixty-six teachers were hired to staff the remediation component across the five sites. Teaching assignments by subject area (reading, math, modules, CAI), effective student/teacher ratios (e.g., class size and the utilization of teacher aides), location of CAI in the classroom or in a lab setting and the number of Cohort I returnees influenced the number of remediation teachers required in each site. In San Diego, where separate teachers were hired for reading and math and class sizes were small, the greatest number of teachers were employed--22. In Fresno--where separate teacher assignments were made for reading, math and modules and one instructor staffed the math CAI lab--15 teachers (eight for Cohort II youth and seven for Cohort I) were employed.

Unlike the structure in Fresno and San Diego, remediation teachers in Boston, Portland and Seattle were responsible for teaching both reading and math. Boston hired 11 remediation teachers for reading and math and one instructor at each university to staff the CAI labs. Boston's low Cohort I return rate, influenced the number of teachers required there. Portland's high number of Cohort I returnees resulted in the assignment of six teachers to each cohort while one teacher taught both groups for a total of 13 reading and math teachers. In addition, two CAI teachers were hired in Portland. Seattle hired the fewest teachers--10. This low number was the result of three factors: teachers taught in both morning and afternoon sessions, the same teacher was used for both reading and math instruction, and no separate CAI instructors were hired.

A somewhat ethnically and racially diverse group of teachers was found, but the proportion of whites was still much greater among the teachers than among the students. The racial/ethnic composition of remediation teachers across sites was as follows: 61 percent white, 27 percent black, 8 percent Hispanic, 3 percent Asian and 1 percent from other racial/ethnic groups. Non-whites accounted for only 39 percent of the teachers and more than 85 percent of the students.

The vast majority (81%) of the teachers hired to provide STEP remediation taught during the regular school year. Most of these remediation teachers taught at the middle/junior high and senior high school levels (grades 7 to 12). Thirty-three percent of them taught English or reading; 16 percent taught math; 2 percent taught health or sex education; and 31 percent taught some other academic subject in their current jobs. Of the remaining remediation teachers, 5 percent were school counselors, 3 percent were school administrators and 11 percent had some other profession.

In addition to general teaching experience, 95 percent of the teachers hired were experienced in working with special needs populations of some kind. Experience with the instructional methods varied, however. While 85 percent of the remediation teachers reported experience with individualized, skills-based instruction, less than half (47 %) had experience with computer-assisted instruction.

The teaching philosophy of most of the teachers was fairly similar to that embedded in the STEP program. When asked how compatible the STEP teaching philosophy was with their own, 57 percent said it was quite close, and 12 percent that it was completely compatible. Twenty-eight percent of the respondents said the philosophy was only somewhat compatible with their own, while 3 percent reported that it was totally incompatible. Overall, teachers liked the module approach to instruction--52 percent liked the modules very much and 48 percent liked them somewhat. Less than 1 percent reported not liking the module approach to instruction.

Teacher Training

Training of STEP teachers improved compared to the first summer. As mentioned earlier, P/PV consultants, lead teachers and other members of the curriculum development team worked together to develop training materials for STEP teachers. Teacher training was provided earlier, and more time was allotted for this activity across sites. All sites provided between three and four days of training for remediation teachers in advance of the start of the summer program.

Assessments by on-site coordinators indicated that teacher training in Boston, Fresno, San Diego and Seattle was executed as planned. In Portland, problems with inadequate enrollment diverted the attention of program operators from training. Across sites, however only 68 percent of the remediation teachers themselves reported that training was adequate; 32 percent felt it was less than adequate. Post-summer discussions with site coordinators and lead teachers identified specific areas for improvement that will be addressed during training for the summer of 1987.

In addition to formal teacher training, some sites provided regular in-service sessions with teachers during the program. In Boston, weekly two-hour sessions were held after the program started, Seattle held hour-long sessions two days weekly, and San Diego provided one mid-program in-service session.

Teacher Aides

Each site hired a number of teacher aides. Their training and utilization were the responsibility of the sites. The backgrounds of the aides hired varied across sites, as did the manner in which they were utilized. Aides' backgrounds ranged from experienced classroom assistants or substitute teachers in four sites to older SYETP youth in Fresno. In some sites, aides actually taught small groups or served as CAI assistants; but in other sites, aides performed mostly clerical and classroom management tasks, such as xeroxing and taking attendance. In Year Three, training for teacher aides will be mandated by P/PV to ensure their most appropriate use.

Computer-Assisted Instruction

Computer-assisted instruction (CAI) is another element of the remediation component that experienced some fundamental changes during Year Two. These changes were based on the experience of the first summer and resulted in better implementation success for CAI in four of the five sites. In all sites, however, there was room for improvement in one or a combination of the following: the sufficiency of hardware, the relevance and comprehensiveness of software, integration of CAI material with classroom instruction, and interaction of teachers with students during CAI.

During Year Two, we found that CAI could be successfully integrated into the remediation component in two ways: in the classroom or in a separate computer lab setting. The specific implementation approach used in each site is described below.

In Boston, a separate CAI classroom setting was used. To improve their implementation of CAI over Year One, the number of hours of access to CAI was increased to three hours each week and CAI experts were added to the staff to conduct CAI activities at each university site. However, CAI in Boston was focused on just math during Year Two. This focus in Boston resulted from their feeling that module activities covered mostly reading (further expanded by Boston's Literature Connection module) and the better availability of math software.

During the first summer, Fresno STEP used a separate computer lab for both reading and math. In Year Two, reading-related CAI was more integrated with the curriculum by placing computers in the skills reinforcement classrooms. Math CAI was again provided in a computer lab set up for that purpose. Fresno used software developed by the San Diego program for providing math CAI. Availability of sufficient copies of this software was the deciding factor in structuring the math CAI component separately. However, teachers accompanied their students to CAI and math CAI was felt to be well-integrated.

Portland improved its CAI component in Year Two by moving the computers out of the classrooms, where they were not effectively utilized during the first summer, and into a computer lab setting where a CAI expert was hired to implement this area of the remediation component. During the first year, regular classroom teachers, who were not in general very familiar with computers, were responsible for providing CAI in Portland.

San Diego's implementation of CAI remained the same for the most part in 1986. At some of its five sites, the computers were in the classroom, while other San Diego sites used a computer room(s) separate from the classroom. In both cases, reading and math teachers were responsible for CAI activities. However, unlike the Year One experience, a sufficient number of computers were available across sites for use by individual students during the CAI period.

CAI in Seattle experienced the most problems in Year Two. A computer lab with 15 computers was staffed by two work/study aides rather than teachers. The lab contained insufficient computers for the number of students and was in poor condition in terms of machine age and equipment maintenance. Furthermore, teachers in Seattle did not have a background in CAI, and no training was provided to them in advance of program start-up. One mid-program session was provided, but was inadequate both in length and content.

Rates of Participation

The unique requirements of a research demonstration necessitated some changes in regular SYETP operational procedures. One of the major changes was in the payroll system, where classroom hours are documented separately from work hours. Setting up systems to record these hours accurately proved to be a challenge in the first year of the demonstration. While the system improved in the second summer, some sites continued to experience difficulty in maintaining accurate records.

Treatment youth were offered an average of 99.0 hours of remediation, including breaks, across sites. The highest number of remediation hours, 112.5, was offered by Boston; Seattle offered 101.5 hours; followed by Fresno, San Diego and Portland which offered 99.0, 96.0 and 94.5 hours respectively.

Total available hours in the classroom were higher, however, when life skills classes were considered. Payroll hours, the source of participation data, combine remediation and LSO hours; work hours are counted separately. The addition of LSO hours to the total revises the maximum available classroom hours as follows: Boston, 133.5; Fresno, 117.0; Portland, 112.5; San Diego, 120.0; and Seattle, 122.5.

Across all sites, youth from Cohort I attended class for an average of 99.0 hours or for 82 percent of the hours available to them. San Diego's youth had the highest rate of participation--89 percent of available hours--while youth in Portland attended class for the lowest proportion of available hours--73 percent.

The classroom participation rates for Cohort II youth were quite comparable to those of the Cohort I returnees noted above. Across all sites, Cohort II youth participated in classroom activity for an average of 83 percent of available hours. The range of participation hours across all sites went from 73 percent in Portland to 91 percent in Fresno. Table IV.5 provides data on mean levels of participation in classroom and work experience for both cohorts of youth.

The classroom participation rates in Portland and Seattle are substantially below that of the other sites. Seattle reported low attendance and observations by P/PV staff confirmed this. However, observations and reports about attendance in Portland suggested that attendance was on a par with other STEP sites. Based on those reports and observations, it is believed that poor payroll record-keeping of hours in Portland may explain some of the discrepancy. Payroll record-keeping problems were particularly severe in Portland where a new computer system was installed.

Planned Changes

Areas identified for improvement for the next summer include better integration of the components, earlier delivery of modules to sites to facilitate program logistics, better utilization of the silent sustained reading feature, better software for use in CAI and better teacher training. Training for aides will be mandated by P/PV.

For Year Three, P/PV will develop an additional five modules in the winter and provide them to sites for comments. (These modules will be in the jobs and life skills areas.) In addition, P/PV will develop a detailed outline for remediation teacher training. Finally, P/PV's remediation consultants will assist sites directly by providing key portions of the training.

THE LIFE SKILLS AND OPPORTUNITIES (LSO) COMPONENT

Life Skills and Opportunities (LSO) is an educational program designed to motivate youth to delay parenthood until they achieve their educational and career goals. The curriculum emphasizes the serious economic, health and social risks of early and unprotected sexual activity. Abstinence is advocated throughout the curriculum; however, in view of the fact that a substantial proportion--42 percent in 1986--of participants reported that

Table IV.5

MEAN HOURS (PERCENT) OF PARTICIPATION IN CLASSROOM AND WORK EXPERIENCE COMPONENTS

	Returning Cohort I Youth			Cohort II Youth		
	Classroom	Work Experience		Classroom	Work Experience	
Site	Treatments	Treatments	Controls	Treatments	Treatments	Controls
All Sites	99 (82%)	105 (80%)	244 (83%)	100 (83%)	107 (81%)	206 (84%)
Boston	115 (86%)	76 (87%)	na	114 (85%)	73 (84%)	166 (93%)
Fresno	103 (88%)	117 (81%)[a]	na	106 (91%)	133 (92%)	244 (91%)
Portland[b]	82 (73%)	83 (62%)	na	82 (73%)	86 (65%)	166 (70%)
San Diego	107 (89%)	138 (79%)	244 (83%)	106 (88%)	133 (76%)	240 (82%)
Seattle	92 (75%)	112 (91%)	na	92 (75%)	116 (95%)	215 (88%)

---

[a] While treatment youth in both cohorts were offered the additional hours of work, the Cohort II group was allowed to work more hours because of poor record-keeping by the caseload coordinator assigned to that group.

[b] The reported number of participation hours in Portland probably understates the true number of participation hours because Portland experienced difficulty with their payroll record system.

na = not applicable

they were sexually active prior to entering to STEP, information on contraception is also provided.

The curriculum comprises two volumes: Volume I is used during the first of the two summers; Volume II for returning participants. Thus, Cohort I received instruction from Volume II during the 1986 summer while Cohort II used Volume I.

The curriculum is designed to engage youth actively in the learning process and employs a variety of teaching techniques to challenge and motivate participants. The techniques include: role-playing--participants are called upon to act out hypothetical situations; a variety of games in which they compete among themselves to answer questions; films; discussions; and small group interactions.

LSO instructors are referred to as "leaders," whose primary role is to prompt and direct discussion among participants rather than to "teach" in the traditional classroom fashion. Leader "lecture time," is limited to not more than 15 minutes in each 90-minute session, and is used to introduce the topic for each session, explain the learning activities and summarize the discussion at the session's conclusion. The only exceptions to this format are the two sessions in Volume I that deal with basic facts about sexuality and contraception; a substantial portion of these sessions is presented by the leader. In all the other sessions, leader time is given to stimulating response from participants and helping them to carry out exercises.

STEP sites followed the same LSO schedule in 1986 as they had in 1985. All offered two 90-minute sessions each week. In every site except Boston, LSO classes were incorporated into the daily remediation schedule. In Boston, participants attended LSO class two afternoons a week following three hours of morning remediation classes. (As discussed in Chapter VI, Boston's scheduling pattern may have been a factor in the small impacts on contraceptive knowledge that were observed in that site.)

Curriculum Content

A typical 90-minute session (in both volumes) consists of three or four activities: a brief introduction of the session's topic by the leader (five minutes); a small-group activity in which several teams of three to four participants, for example, perform a role-play or develop a solution/response to a problem (40 minutes); a film and follow-up discussion (40 minutes); a concluding summary by the leader (five minutes).

Broadly, Volume I introduces participants to the basic skills and knowledge they will need to make responsible choices about their sexual behavior. Specific topics include:

- The role of personal values and attitudes in shaping life goals and behavior;
- The negative and limiting effects of sex-role stereotyping and how to change them;
- How to make responsible decisions and overcome pressure to change them;
- Handling the pressures of dating relationships, understanding sexual feelings and managing them responsibly;
- The personal and financial responsibilities and costs of parenthood;
- Key facts about sexuality, including anatomy, contraception and sexually transmitted diseases; and,
- Methods of contraception, including abstinence.

Toward the end of the program, first-year participants also make an informational visit to a comprehensive health facility. This feature is included because many youth are not aware of the availability or location of adolescent health services. Acquainting youth directly with a local facility, where they meet staff and learn about the range of available services, helps to alleviate the reluctance and anxiety often associated with the use of medical facilities.

Volume II reinforces the material presented in Volume I and introduces some new material. Volume II reinforces first-summer explorations of decision-making, relationships, facts about sexuality, contraception and the risks of unprotected sex. The new areas in Volume II examine:

- The importance of standing up for oneself;
- Appropriate work attitudes and behavior;
- How to set career goals and plan for the future;
- Maintaining good health;
- The risks of substance abuse; and,
- How to cope with a variety of stressful situations--particularly those that arise in the family and other close relationships.

Implementation of the LSO Component

Overall, LSO was the most smoothly implemented component of the 1986 summer, as it was in the previous year. The curriculum is a complete, fully-developed package that provides the instructor with all the information s/he will need to conduct each session, as well as with clear instructions on each session's objectives, materials, and the time to be allocated to each activity (see sample). The fact that students have considerable interest in the subject also contributes to LSO's successful implementation.

Not unexpectedly, Volume I implementation was the most successful in Year Two. Its contents had been refined by two summers of experience (in the 1984 pilot and the demonstration's Year One), and one-third of the teachers had taught the curriculum at least once. In addition, the Cohort II participants who received Volume I instruction appear to have been more attentive than the returning Cohort I teenagers who received instruction from Volume II.

The field visit to a neighborhood health clinic, part of the Volume I curriculum, was reported by teachers to be an important and informative event for STEP participants, most of whom were unfamiliar with their clinic's location and services. About half the visits were well-planned, with clinic staff making presentations specifically designed for the STEP group and eliciting questions that indicated serious interest on the students' part. Other visits, however, did not involve sufficient advance planning to engage the interest of youth beyond learning that the clinic was there and ready to help them.

Volume II implementation exhibited the problems to be expected with a new curriculum. Teachers had to become familiar with the new material and sequence; many indicated that they had difficulty covering all of the activities indicated for each session; and more teachers of Volume II reported problems holding the interest of the participants, particularly in the three sessions dealing with career and job behavior. A number also reported difficulty with small group management.

LSO Instructor Training

Consistent with the other instructional component of STEP, P/PV required LSO teachers to be trained in presenting the curriculum. LSO instructors attended an intensive, two-day training session delivered by the curriculum developers. Two sessions were held for the 1986 instructors--one in Seattle attended by the four West Coast sites and a second session in Boston. As in the first summer, instructors' response to the training was positive in Year Two. While instructors almost uniformly would have preferred longer sessions, the training was very well received. The training appeared to be effective in familiarizing teachers with curriculum content and instructional techniques. It also generated a high level of enthusiasm.

LSO Teacher Characteristics

Thirty teachers provided LSO instruction in the five sites. More than half of them were currently employed as teachers during the school year. Thirty-one percent taught regular academic subjects; 21 percent taught health or sex education courses. Of the remainder, 17 percent were counselors (either in schools or social agencies), 7 percent were in school administration and 24 percent held some other job. A high proportion, 86 percent, had previous experience with special needs populations. In contrast to their remediation counterparts, a majority of LSO teachers (53%) were male. LSO teachers also reflected more closely the racial mix of participants--half of the staff were black, 43 percent were white and 7 percent were Hispanic.

As will be discussed in Chapter VI, none of the objective characteristics that were measured--sex, race, teaching experience, occupation, experience with special needs populations--seemed to have had any affect on life skills outcomes. However, our observations of LSO teachers indicated wide variation across sites in teacher performance. Based on these observations, we have identified several features that we believe are critical to effective LSO instruction:

- Strong belief in and commitment to the importance of the curriculum;
- Comfort with the material and with the youth;
- Ability to make teenagers feel comfortable with the topics;
- Experience leading group activities; and,
- A sense of humor.

Planned Changes

Although the conduct of small group activities is discussed in both volumes of the curriculum, last summer's teacher feedback and site observations indicated that more specific direction is needed and will be provided.

A "key concepts" sections will be added at the beginning of each session to highlight for leaders the critical issues that should be covered.

The lead agency in each site will be specifically charged with scheduling the first-summer clinic visits to ensure that they are more sufficiently planned than in the previous summer.

Portions of Volume II are being revised. Activities in some sessions will be streamlined or, in some cases, eliminated. The material on jobs and careers will be reworked and consolidated into two sessions. The revisions will reduce possible overlap with material that is covered in the remediation modules or in the orientation to work experience that participants received at the beginning of the summer.

THE WORK EXPERIENCE COMPONENT

The provision of practical work experience is a major feature of the STEP intervention and of the demonstration. Program participants are offered a combination of work and remediation, while control group members, whose experience is compared with that of the treatment group's, are offered only work.

Component Description

The local agencies administering SYETP in each site are responsible for operating the work experience component. They secure part-time jobs for all the treatment youth, full-time jobs for all the control youth, and hold an orientation session for both groups of youth (as is routine in SYETP). The jobs obtained for the youth were typical, for the most part, of SYETP jobs obtained for 14- and 15-year-old youth. These included positions as custodial or maintenance aides, recreation aides, clerical helpers and day care aides. Most sites were able to find some atypical job assignments--for example, working at a boat house doing maintenance and teaching sailing techniques, as a news assistant at a university radio station, in an animal care laboratory, at a naval hospital and air station, and in a theatrical production.

The amount of work experience provided by STEP was prescribed by the demonstration design. Treatment youth were offered approximately 80 hours of part-time work in combination with approximately 120 hours of remediation. Control group youth were offered approximately 200 hours of employment. Thus, the number of hours spent in a structured activity during the summer should be comparable for treatments and controls. Treatment youth worked between three and three and one-half hours each afternoon (except in Seattle where half of Cohort II and all of Cohort I worked in the morning). Control youth generally worked between six and seven hours a day. In Boston, however, control youth worked only 25 hours per week.

Implementation Issues

Youth in regular SYETP work full time. Obtaining part-time jobs for STEP youth was new for SYETP staff during Year One and proved to be difficult. However, based on the first year's experience, SYETP staff encountered less difficulty in this area in Year Two and were fairly successful in securing part-time slots. One strategy used for securing these slots was linking full-time slots with part-time slots. Employers were provided with a certain number of full-time (control group) workers, in exchange for taking a certain number of part-time workers. At first employers were somewhat resistant to part-time work. However, work-site observations conducted by the site coordinators disclosed that the work performed by the STEP youth was beneficial to both the youth and the employer. Employers who provided part-time jobs the first summer were quite willing to do so again.

Participation rates of treatment and control youth differed for a variety of reasons. For example, program rules specify that treatment group members could not work if they did not attend their remediation and LSO classes that day. Program participants, thus, might work fewer hours than controls if they skipped classes more often than the control group skipped work. However, in Table IV.5, we find that participation rates among controls were only slightly higher, on average, than they were for treatments. Overall, control group youth attended 84 percent of the offered work hours, while the treatment group youth attended 81 percent.

The research design specifies that treatments and controls are offered approximately the same number of hours of structured activity (be it employment or school). The only site where the hours greatly differed between groups was Boston. Contrary to their original plans, Boston offered controls fewer hours than it did to treatments. Treatment youth were offered a total of 220.5 hours, while control youth were offered a total of 178 hours.

The goal of the program is to find jobs, to the extent possible, that match the interests of the youth. Accomplishing the match for the incoming cohort is difficult since treatment or control status is revealed only after youth are pre-tested. Therefore, SYETP staff cannot know whether to develop a part-time or full-time job for a particular youth. To ensure that everyone has a job, jobs are identified before experimental status is determined. After random assignment of the youth, the pre-selected jobs are allocated to best match the youths' interests. A better match is possible for returning youth who are known to the SYETP staff. In fact, better jobs were generally obtained for returning youth because of this ability to match without concern about random assignment. In addition, because returning youth are older and have some work experience, the STEP program policy is to offer better jobs to second-year participants.

THE SCHOOL-YEAR SUPPORT (SYS) COMPONENT

The 1985-86 school year was the demonstration's first opportunity to implement a school-year support component. All treatment youth from the first cohort were eligible to participate in SYS activities regardless of their level of participation in the summer program.

The objectives of SYS as defined in the first year were threefold: 1) to encourage STEP participants to continue their achievements of the summer; 2) to provide a continuing means of keeping track of participants; and 3) to link participants with tutorial, health, counseling and other support services. Each site developed a school-year support plan consistent with these objectives outlining the design, management and staffing of their program. P/PV required each site to designate a coordinator to assume responsibility for the SYS program and to offer two levels of services/activities to participating youth--group activities and individualized contact. Parental involvement in SYS was encouraged. Sites received modest grants to carry out school-year support and were responsible for providing P/PV with monthly participant contact reports.

Implementation

Four of the five sites submitted plans that were approved by P/PV. San Diego declined to implement a school-year support component, and none of the four plans that were approved were carried out as designed. As a result, the functioning of SYS as a component of STEP in 1985-86 cannot be judged since, in practice, it was not very strong for Cohort I. It remains to be seen whether or not it is possible for STEP youth to sustain their group identity during the school year when they are dispersed throughout the site.

Except in Fresno, the entity responsible for operating the summer program in each site was also responsible for SYS; most sites assigned part-time coordinators who were also responsible for higher-priority tasks. Each of the four sites planned and held a series of large group events--athletic, social and educational--but they were poorly attended. Each of the sites assigned counselors to meet periodically with STEP youth--university students, except in Portland, which used school staff--but these individual contacts were not made regularly.

Fresno's SYS program included incentive programs, a monthly newsletter, an awards banquet, field trips, an active advisory group of parents, and better attendance than in other sites; but even so, the program did not match either the site's or P/PV's expectations for continuing contact with STEP youth.

Program Redesign

The disappointing results of the first year made it necessary to re-assess the overall program structure and management of the SYS component. P/PV made structural and managerial changes for the 1986-87 program and has increased its oversight of this component by designating a staff member to work with the SYS staff at each of the sites.

The objectives of the SYS program have been re-stated for implementation in the second year. Specifically, SYS will:

1. Encourage STEP participants to return for a second summer of STEP treatment;
2. Provide a continuing means of keeping track of participants;
3. Motivate youth to stay in school by identifying barriers to school retention and providing referral to needed services (e.g. tutoring, counseling, medical); and
4. Increase the involvement of the school system in STEP.

While the basic requirements for SYS remain unchanged, they have been more strictly defined for the second year as follows:

- A coordinator will manage the SYS program in each site and will work on SYS activities a minimum of 20 hours each week. The coordinator can be a local staff member or consultant, or a consultant to P/PV. (P/PV participated in the selection of SYS coordinators for 1986-87.)
- Mentors/counselors will be experienced in such relevant areas as youth counseling, social work, education, etc. They will be trained for their SYS role by the coordinator and must work at least 15 hours each week with youth. Caseloads will be between 15-20 students.
- Support and contact activities will occur in both large group settings and individualized meetings.
- Parents will be regularly informed about SYS activities and encouraged to help with group events.

- The school system will be encouraged to participate in the planning and implementation of SYS activities.

A redesigned monthly reporting format has also been instituted. Coordinators submit monthly contact forms that identify the nature and level of contact with youth. In addition, they prepare descriptive monthly reports identifying key issues and problem areas. All five sites are operating an SYS component during the 1986-87 school year.

SUMMARY

Identifying and recruiting youth to participate in STEP was greatly improved in the second year of the demonstration, and efforts to attract the required number of youth to the demonstration were successful. Nevertheless, these activities continued to require a heavy investment of time and effort, particularly in re-enrolling Cohort I treatment youth for their second summer. Targeted recruitment, personal contacts with youth and active involvement by school personnel were found to be necessary. Since the School Year Support component did not function as planned, its role in tracking and encouraging youth to continue in STEP did not materialize.

Testing procedures were revised to ensure that both treatment and control youth were exposed to precisely the same conditions, thereby protecting the integrity of the net differences in their performance.

Planning for the introduction of a new remediation curriculum, Practical Academics, was initiated well in advance of the summer program; lead teachers were designated and began work early in the spring; and STEP teachers were hired and trained in time to prepare for the summer. For this component, some start-up time was required in the second year of the demonstration, as teachers familiarized themselves with Practical Academics' modules. Changes were also made to increase the success of computer assisted instruction. Among changes planned for Year Three are better utilization of the new sustained silent reading feature and improved teacher training.

Participation rates continued to be good. Cohort I youth participated in classroom activity for an average of 82 percent of available hours, Cohort II youth for 83 percent.

The LSO curriculum was, as in Year One, smoothly implemented and well received by teachers and STEP participants. Expected start-up problems attending the introduction of Volume II for use in the second summer of participation were minimal and will be corrected in Year Three.

The third component, work experience, continued to provide low-skill, part-time SYETP jobs for the 14 and 15 year olds in STEP. Sites did manage, however, to upgrade a number of jobs for returning Cohort I youth and to reduce the number of custodial and maintenance positions.

The School Year Support component, which did not function as planned, has been substantially strengthened for school year 1986-87. Its operation in Year One cannot be judged since it involved, in practice, very few Cohort I youth. All five sites are now operating an SYS program in compliance with the design, and its assessment will be included in the next research report.

V. PARTICIPANT BASELINE CHARACTERISTICS

This chapter describes the two cohorts of the STEP sample--youth who began participating in the summer of 1985 (Cohort I) and youth who joined STEP in the summer of 1986 (Cohort II). The characteristics of Cohort I were described in detail in an earlier report (Branch et al., 1986). This chapter, therefore, focuses on the baseline characteristics of Cohort II, but highlights differences in the attributes of the two cohorts. Differences in the characteristics of treatment and control youth enrolled in the Cohort II sample are noted if they reach statistical significance.

A description of returning members of the Cohort I treatment group is also provided in this chapter. The returnees, as this group is frequently called, will be compared to non-returning members of the treatment group in the Cohort I sample. The chapter concludes with a summary of the control youth in San Diego who returned for a second summer of work experience.

The data that describe the Cohort II sample are based on questionnaires and MAT pre-tests in reading and math administered at the 1986 baseline, while tables on returning youth are based on 1985 baseline data. Tables presenting characteristics of the aggregate sample of 2,431 youth in the learning analysis sample can be found in Appendix C. Throughout the chapter, characteristics of the aggregate sample are mentioned only when they vary substantially from those of the second cohort.

THE COHORT II SAMPLE

As discussed earlier, 1,635 youth were randomized to the Cohort II sample of STEP. However, not all of these youth actually participated in STEP. In this section, therefore, we describe only the 1,268 youth who make up the learning analysis sample for Cohort II. The learning analysis sample includes all Cohort II youth with pre-test and post-test MAT scores and valid responses to all of the variables included in the multivariate learning analysis.[1]

---

[1] The variables included in the learning analysis regression models are: treatment status, sex, race/ethnicity, baseline measures of family size, birth order, grade in previous school year, grade retention, absenteeism, expectations of post-high school education, problem with English, enjoyment of reading and math, reading and math MAT scaled scores and demonstration site dummy variables.

Since the integrity of the impact estimates depends critically on the kind of attrition observed from the randomization to analysis samples, we conducted statistical tests to determine whether differential attrition occurred between the two samples. It is important that the sample generating the estimated program impact be representative of the original sample of youth. If attrition differs systematically between treatment and control youth, the estimates of program impact may be biased. Further, if substantial changes in the composition of the sample with respect to key characteristics occur, this can affect the generalizability of the results. For example, if a disproportionate percentage of one subgroup (such as males or Hispanics) in the randomization sample is lost to the analysis sample, program impacts for the affected subgroups may be over- or under-estimated.

Statistical tests indicate that there are no significant differences between treatment and controls with respect to the sex, age and race/ethnicity distributions in the randomization and learning analysis samples. Also, treatment status was not found to affect the probability of being in the learning analysis sample.[2] Since there is no evidence of differential attrition, we are more confident that the impacts presented in this report represent unbiased and consistent estimates of the treatment effect on the analysis samples.

Table V.1 presents key demographic characteristics for the sample of youth who were randomized to Cohort II and the sample available for learning analyses.

Basic Demographic and Educational Ability Characteristics

The Cohort II sample of STEP represents a diverse group of youth with varied backgrounds but similar educational needs. The 1,268 youth who comprise the learning analysis sample for this cohort are a group at risk of dropping out of high school based on their academic performance histories and economic backgrounds. The Cohort II sample met the requirements for participation in STEP on all three criteria--age, economic status (all are JTPA-eligible) and educational level. The basic demographic and

---

[2] Estimates of the probability of being in the learning analysis sample were derived using logit analysis. Similar attrition analyses were completed for the aggregate sample. The aggregate results also indicate that treatment status did not affect the probability of being in the learning analysis sample.

Table V.1

SELECTED CHARACTERISTICS OF THE RANDOMIZATION AND LEARNING ANALYSIS SAMPLES FOR COHORT II

Characteristic	Randomized Sample Treatment	Randomized Sample Control	Randomized Sample Total	Learning Sample Treatment	Learning Sample Control	Learning Sample Total
Sex						
Male	47%	53%	50%	47%	51%	49%
Female	53	47	50	53	49	51
Age						
14 and Under	54%	55%	54%	53%	56%	55%
15 and Over	46	45	46	47	44	45
Race/Ethnicity						
Black	46%	46%	46%	48%	48%	48%
Hispanic	19	18	19	18	18	18
Asian	18	18	18	19	18	19
White & Other	17	18	17	15	16	15

NOTE: Of the 1,635 youth who were randomized, 119 were never pre-tested. For the learning analysis sample an additional 123 youth were omitted because they were missing post-test scores. The remainder were omitted because of non-response on at least one variable in the learning analysis equation, leaving 1,268 youth.

educational ability characteristics of these youth are discussed below and presented in Table V.2.

Age

Fifty-five percent of the Cohort II sample were 14 years old when they entered the program. The excess of younger youth is consistent in all STEP sites except Seattle, where 52 percent were 15 years of age at entry. It is not surprising that sites enrolled more 14-year-olds than 15-year-olds for two reasons. First, considering the usual difficulty encountered by younger, inexperienced teens in the labor market and the fact that SYETP traditionally serves very few 14-year-olds, STEP provides a unique opportunity for this age group. Second, because STEP youth are recruited from the same school systems each year, it is likely that a substantial proportion of the available pool of 15-year-olds were enrolled in the first cohort of the sample.

Sex

STEP appears to be equally attractive to males and females. Across all sites, for both Cohort II and the aggregate sample, approximately half of Cohort II is male and half is female.

Racial/Ethnic Composition

Cohort II is a more diverse group racially and ethnically than Cohort I. While more blacks were served than any other racial/ethnic group, overall participation by blacks dropped from 59 percent to 48 percent for Cohort II. More Hispanics, Asians and youth from other racial groups were served in Cohort II than in the previous cohort. Asian participation went from 14 to 19 percent, participation of Hispanics went from 16 to 18 percent and participation of other groups increased from 1 to 6 percent. The participation of whites dropped one percentage point during the second summer--from 10 to 9 percent.

As was true in the first summer, racial/ethnic composition varied by site. The largest variations from Cohort I to II were evident in the Fresno and Boston sites. For Cohort I, less than 1 percent of the Fresno sample was Asian compared to 14 percent in Cohort II. Similarly, whites and others made up only 5 percent of the Fresno sample during the first summer, but these groups combined for a total representation of 12 percent during 1986. These changes in Fresno were the result of organized efforts to enroll a sample more fully representative of Fresno's JTPA-eligible population. In Boston, the proportion of Hispanics in STEP increased by eight percentage points from Cohort I (13% versus 21%).

Table V.2

DEMOGRAPHIC AND EDUCATIONAL ABILITY CHARACTERISTICS OF COHORT II

Characteristic	All Sites	Boston	Fresno	Portland	San Diego	Seattle
Age						
14 Years Old at Entry	55%	58%	57%	58%	53%	48%
15 Years Old at Entry	45	42	43	42	47	52
Sex						
Male	49%	45%	54%	50%	47%	48%
Female	51	55	46	50	53	52
Race/Ethnicity						
Asian	19%	7%	14%	15%	33%	26%
Black	48	59	30	50	42	58
Hispanic	18	21	44	3	18	2
White	9	8	8	23	2	7
Other	6	5	4	10	5	7
Educational Ability						
Mean MAT Reading Score at Entry (Grade Equivalent)[a]	722 (6.0)	704** (5.1)	727 (6.3)	726 (6.3)	731 (6.5)	724 (6.2)
Mean MAT Math Score at Entry (Grade Equivalent)[a]	703 (6.7)	679 (6.1)	710* (7.0)	701 (6.7)	733 (7.8)	694 (6.5)
Sample size	1268	262	256	246	258	246

***Treatment-control difference is significantly different from zero at the 0.01 level.

**Treatment-control difference is significantly different from zero at the 0.05 level.

*Treatment-control difference is significantly different from zero at the 0.10 level.

[a] The score differed among sites at a 0.01 level.

Educational Ability

As discussed in the chapter on program implementation, sites used different standardized tests for screening potential STEP-eligible youth. However, since the Metropolitan Achievement Test (MAT) was administered to all youth at baseline, there is a uniform basis for assessing the academic ability levels of youth at program entry.

All youth scored at the 6.0 grade equivalent level in reading, at least two grades below where they should be based on their age. As Table V.2 shows, youth selected for STEP scored within a range of 0.3 of a grade equivalent on reading in four of the five sites (6.2 to 6.5). In Boston, however, youth were considerably more disadvantaged, scoring almost a full grade equivalent below the next lowest site--at the 5.1 grade level. (As will be discussed later, the fact that Boston enrolled a substantially higher proportion of eighth graders likely accounts for some of this difference in entry-level scores.)

Overall, youth scored at the 6.7 grade level in math, at least a year and a half behind where they should have been in school. Math scores varied significantly among sites. Youth in Boston again had the lowest scores (6.1 grade level), while scores in San Diego were highest (7.8 grade level).

Across all sites, Cohort II youth look similar to Cohort I in terms of both their reading and math ability levels. In reading scores, Cohort II tested at the 6.0 grade level, while Cohort I tested at the 5.9 grade level. For math, Cohorts I and II scored at the 6.5 and 6.7 grade levels, respectively.

While there were essentially no differences in the tested levels between cohorts across all sites, site-by-site differences between the two cohorts were quite large. For reading, entry level scores for Cohort II youth in Boston were significantly below those of Cohort I (704 versus 729). Conversely, scores in Fresno and Seattle were significantly higher for Cohort II. In Fresno, Cohort II scored at the 6.3 grade level, compared to scores at the 5.6 grade level for Cohort I. In Seattle, Cohort II youth scored at the 6.2 grade level compared to scores at the 5.0 grade level for Cohort I.

Similarly, math scores differed greatly within sites between cohorts. Cohort II youth in Seattle and Fresno scored significantly higher than their Cohort I counterparts at baseline. Seattle's Cohort II youth received average scores of 694 in math (6.5 grade level) compared to an average score of 667 (5.8 grade level) for Cohort I. In Fresno, Cohort II scored at the 7.0 grade level, while Cohort I scored only at the 6.5 grade level. (In Boston and Portland, differences between the two cohorts in math were significant at the 0.10 level.)

Despite the slightly higher entry scores by Cohort II, STEP sites were again able to enroll youth who are educationally disadvantaged and likely to drop out. Overall academic test scores indicate that Cohort II youth are performing at least one and one half grade levels below their expected grades of enrollment, which are eighth through 10th grades.

No significant differences between treatments and controls were found with respect to baseline scores overall. In Boston, however, controls scored significantly higher--by 15 points--than treatments in reading at pre-test. Tables displaying entry level test scores by age, sex and racial/ethnic group for the Cohort II sample can be found in Appendix C.

Additional Demographic and Economic Characteristics

Language Barriers

Forty percent of the youth in Cohort II have spoken or currently speak another language in addition to English. (See Table V.3.) Of those youth who report speaking a language other than English; 23 percent reported having problems understanding teachers who speak English; 28 percent reported having problems communicating in English; and 29 percent reported problems reading English. Therefore, a substantial minority of these youth are testing poorly because of their limited command of the English language.

Family Composition

Female-headed households are particularly prone to be economically disadvantaged. Approximately half of Cohort II youth come from families headed by a female. The variation was also statistically significant across the sites. In particular, Boston enrolled the most students from these families--65 percent; whereas Cohort II youth in Fresno had the lowest reports of female-headed households--40 percent. These patterns hold for the aggregate sample as well, where exactly half of the sample are from households headed by a female.

The mean household size for Cohort II youth is 5.16, nearly equal to that of the aggregate sample. Twelve Cohort II youth have children of their own, and seven are married.

Employment History

STEP provided the first job opportunity for over two-thirds of the Cohort II sample. Only 29 percent of Cohort II youth had prior work experience. In Cohort I, a larger fraction of the youth had work experience--36 percent. Thus, during the second year of operation, sites enrolled substantially more youth who had no prior work experience. One exception is Seattle where an

Table V.3

ADDITIONAL DEMOGRAPHIC AND ECONOMIC CHARACTERISTICS OF COHORT II

Characteristic	All Sites	Boston	Fresno	Portland	San Diego	Seattle
Percent Who Speak a Language in Addition to English	40%	32%	48%	25%	58%	37%
Of Those Speaking a Language in Addition to English, Percent With						
Problems Understanding English	23%	21%	20%	16%	30%	22%
Problems Speaking English	28	21	29	26	32	30
Problems Reading English	29	21	26	34	30	33
Percent from Female-Headed Households[a]	52%	65%	40%	50%	46%	57%
Mean Household Size	5.16	4.97	5.27	4.75	5.88	4.89
Number Married	7	0	3	1	3	0
Number with Children	12	0	4	1	1	6
Percent Ever Worked	29%	33%	27%	36%	21%	29%
Sample Size	1,268	262	256	246	258	246

[a] The percentage of youth from female-headed households differed significantly across sites at the 0.01 level. The substantially higher rates of enrollment of these youth in Boston and Seattle likely account for this variation.

almost equal proportion of youth from both cohorts had prior work experience.

School Participation and Educational Expectations Characteristics

This section looks at some of the school-related characteristics of Cohort II youth, such as the current grade of enrollment compared to the expected grade of enrollment, a history of being held back in one or more grades, absenteeism and expectations for higher education. Research suggests that these factors are associated with the decision to drop out. The data displaying these characteristics are presented in Table V.4.

Grade of Enrollment

Assuming normal progress, youth ages 14 and 15 are expected to be between the eighth and 10th grades (depending on birth date), and performing at least at those levels. On average, the mean grade of enrollment for Cohort II youth is 8.3. Across sites, 35 percent of STEP youth were enrolled in the 9th grade--the expected grade level for 14-year-olds. The remaining 5 percent of youth were enrolled in the 10th grade. In Boston, where the focus is on serving eighth graders scheduled for transition into a high school environment, 89 percent of Cohort II enrollees are in the eighth grade. Pre-test scores presented earlier indicated that Cohort II youth were performing well below their age and grade-appropriate levels (at the 6.0 grade level in reading and the 6.7 grade level in math), and are in need of an intervention like STEP.

Grade Retention History

A second indicator of educational deficiency and dropout behavior is grade retention history. Nearly one-third (31%) of the Cohort II sample reported having repeated at least one grade.

Absenteeism

STEP youth were asked to indicate the number of times they were absent during the 1985-86 school year from among three categories: 1-10 days, 11-20 days, and more than 20 days. Seventy-three percent of the sample reported absences of between 1 and 10 days, 18 percent reported absences between 11 and 20 days and 9 percent reported absences of more than 20 days.[3] The absentee rates for the aggregate sample mirror those for Cohort II.

---

[3] From the school data for academic year 1985-86, we know that self-reported absenteeism is severely under-reported.

Table V.4

SCHOOL PARTICIPATION AND EDUCATIONAL EXPECTATIONS CHARACTERISTICS OF COHORT II

Characteristic	All Sites	Boston	Fresno	Portland	San Diego	Seattle
Grade in School						
8th Grade or Below	60%	89%	50%	49%	55%	53%
9th Grade	35	10	44	49	37	35
10th Grade	5	1	6	2	7	12
11th Grade	<1	0	0	0	1	<1
Grade Retention History						
Percent Ever Repeating a Grade	31%	47%	35%	23%	21%	30%
Absenteeism						
Percent Absent 1-10 Days in SY 85-86	73%	74%	64%	76%	80%	74%
Percent Absent 11-20 Days in SY 85-86	18	18	22	19	14	15
Percent Absent More than 20 Days in SY 85-86	9	9	14	5	6	11
Educational Expectations						
Percent Planning to Pursue Post-High School Education[a]	72%	63%	75%	76%	74%	74%
Sample Size	1,268	262	256	246	258	246

[a] The difference in educational expectations for youth in Boston is significantly below that of youth in all other sites (at a 0.01 level).

Educational Expectations

Across all sites, 72 percent of Cohort II youth plan to pursue some form of post-high school education. Expectation levels were quite consistent in four of the five sites, where the reports ranged from 74 percent to 76 percent. However, in Boston youth reported significantly lower educational expectations: only 63 percent reported plans to pursue education after high school.

Fertility-Related Characteristics

At baseline, a questionnaire was administered to determine the level of knowledge on fertility-related issues, as well as to collect information on attitudes and behavior with respect to these issues. Because of the sensitive nature of these issues, questionnaires were handled in such a way as to protect the confidentiality of responses. Identification numbers rather than names were used on questionnaires, and students sealed their completed forms in envelopes to insure confidentiality. Nonetheless, non-response rates on fertility-related items were higher than for the other items, but were substantially lower than in the first summer. To indicate items on which the relevant sample size varies due to this non-response, the number of respondents is noted separately on Table V.5.

At baseline, 42 percent of the Cohort II youth who responded reported being sexually experienced. This percentage is significantly lower (by 17%) than that of Cohort I at the same point last summer, perhaps due to the enrollment of a higher percentage of younger youth. Of the 42 percent of Cohort II youth who reported being sexually experienced, 53 percent reported activity in the two months prior to the start of the program. One-third of these recently sexually active youth said they never used any method of birth control during that period.

On average, Cohort II youth who were sexually active at baseline initiated sex two years prior to joining STEP--at the age of 12. Fourteen of the girls had been pregnant, and 16 of the boys had been responsible for a pregnancy.

Cohort II youth have a low level of knowledge on birth control issues. To illustrate, in a series of eight statements designed to test knowledge of the accessibility and availability of various birth control methods, STEP youth answered fewer than half (3.1) of the statements correctly. To test general knowledge about the use and effectiveness of birth control methods, a series of 12 questions was asked. Again, across sites, youth could correctly answer fewer than half of the statements (5 of the 12). No site-by-site differences with respect to knowledge were evident, nor were any significant differences in knowledge or experience found between treatments and controls across sites.

Table V.5

FERTILITY-RELATED CHARACTERISTICS OF COHORT II

Characteristic	All Sites	Boston	Fresno	Portland	San Diego	Seattle
Percent With Prior Sexual Experience	42%	49%	35%	46%	37%	46%
Percent of the Sexually Active Sample Who Were Sexually Active in the Past Two Months	53%	55%	55%	52%	54%	51%
Percent of Recently Active Sample Never Using Contraception in the Past Two Months	33%	34%	39%	32%	42%	18%
Mean Age at Onset of Sexual Activity for the Sexually Active Sample	12.1	11.9	12.6	12.0	12.0	12.1
Number of Females Ever Pregnant	14	2	1	3	0	8
Number of Males Ever Fathered a Child or Responsible for a Pregnancy	16	4	2	1	4	5
Mean Number Correct on Birth Control Availability Questions (8-item scale)[a]	3.1	3.0	3.1	3.2	3.1	3.1

Table V.5 - (cont'd)

FERTILITY-RELATED CHARACTERISTICS OF COHORT II

Characteristic	All Sites	Boston	Fresno	Portland	San Diego	Seattle
Mean Number Correct on General Birth Control Knowledge Questions (12-item scale)[a]	5.0	4.8	5.2	5.2	5.1	5.1
Percent Believing That Both Partners Should Be Responsible for Birth Control	74%	64%	77%	76%	75%	79%
Percent Who Have Taken a Prior Sex Education Course[b]	70%	58%	61%	73%***	88%	66%

NOTE: Across all sites, 1,176 youth answered the sexual experience question: 496 answered the question on sexual activity within the last two months, 288 answered the contraception question, 523 answered the question on the age of sexual initiation, 182 girls and 331 boys answered the question on pregnancy; 1,263 answered the responsibility question and 1,194 answered the question on taking prior sex education courses.

***Treatment-control difference is significantly different from zero at the .01 level.
**Treatment-control difference is significantly different from zero at the .05 level.
*Treatment-control difference is significantly different from zero at the .10 level.

[a] The mean of this variable did not statistically differ across sites.

[b] The level of exposure to sex education by youth differed by site at a 0.01 level of significance.

Inconsistent with the level of knowledge displayed, a surprisingly high proportion (70%) of Cohort II youth reported that they had previously taken a course in sex education. Reports of exposure to sex education varied significantly from site to site, with youth in San Diego indicating the highest exposure level--88 percent--and youth in Boston revealing the lowest--58 percent.

Overall, no significant differences between Cohort II treatments and controls were found with respect to the exposure to sex education. However, in Portland, treatments were significantly more likely to have been exposed to sex education than were controls.[4]

Given the lack of knowledge about birth control issues, the need for curricula addressing these topics is clear. The need is further underscored by the fact that a substantial proportion (42%) of respondents reported previous sexual experience.

THE COHORT I RETURNEES

A key area of interest in any multi-year program is identifying the type of participants who will return after the first year for additional services. This section addresses that issue by looking at the baseline characteristics of the subset of Cohort I treatment youth who returned to STEP in 1986, and compares these with the baseline characteristics of the non-returning group of treatments. (As discussed earlier, only the treatment group was asked to return for a second summer, except in San Diego.)

Of the 564 treatment returnees who were pre-tested, 427 compose the learning analysis sample for the returning Cohort I treatment group across all sites. All data presented in this section come from program questionnaires and MATs in reading and math administered at the 1985 baseline.

Basic Demographic and Educational Ability Characteristics

Age

Across sites, a higher percentage of the younger enrollees returned for the second summer which suggests they may have fewer additional opportunities for summer employment. Fifty-three percent of the returnees were 14 years old at baseline, while only 47 percent of the non-returnees were 14.

---

[4] The difference in exposure level between treatments and controls in Portland is significant at the 0.01 level.

Sex

Across all sites, more females returned for a second summer of STEP than did males. For males, just over 50 percent returned, whereas for females 57 percent returned. Thus, females constituted a larger percentage of the returnees (51 %) than the non-returnees (45 %). Within sites, some differences were also observed. Seattle served more males in the first year, but significantly more females returned to the program in 1986.

Racial/Ethnic Composition

Race/ethnicity was related to the decision to return to STEP for a second summer. The racial/ethnic composition of the returnees compared to that of the non-returnees is statistically significant overall, as well as in Boston, Portland and Seattle. Asians were much more likely to return for a second summer of STEP than were youth from white or other racial/ethnic groups.

Educational Ability

Differences in entry-level academic ability do not appear to influence the decision to return to STEP. The pre-test scores of youth who returned for a second summer were virtually identical to those of the non-returning group. In reading, the returnees scored only one-tenth of a grade level below the non-returnees, while math scores for the two groups were different by only two-tenths.

These basic demographic and educational ability characteristics of the treatment returnees at baseline 1985 are presented in Table V.6. The table also compares the returnees with the non-returnees in the learning analysis sample from Cohort I.

Additional Demographic and Economic Characteristics

Limited English Proficiency

In terms of proficiency with the English language across sites, the returnees were significantly more likely to report limited English proficiency than the non-returnees. (See Table V.7.) The difference is particularly noticeable in Seattle where 20 percent of the returnees reported such problems compared to only 7 percent of the non-returning group.

Family Composition

Across sites, the returnees were similar to the non-returnees in terms of their family composition. Two exceptions were found within STEP sites, however. In San Diego, the returnees came from significantly larger families (5.68 versus 4.93 members),

Table V.6

DEMOGRAPHIC AND EDUCATIONAL ABILITY CHARACTERISTICS OF THE COHORT I RETURNING TREATMENTS

	All Sites		Returnees				
	Non-Returnees	Returnees	Boston	Fresno	Portland	San Diego	Seattle
Age							
14 Years Old at Entry	47%	53%	63%	54%	49%	60%	40%
15 Years Old at Entry	53	47	37	46	51	40	60
Sex							
Male	55%	49%	39%	54%	56%	51%	48%
Female	45	51	61	46	44	49	52
Race/Ethnicity							
Asian	11%	16%	0	0	11%	35%	34%
Black	60	61	81	50	74	45	61
Hispanic	15	16	12	45	4	16	0
Whites & Other	14	7	7	5	11	4	5
Educational Ability							
Mean MAT Reading Score at Entry (Grade Equivalent)	717 (5.8)	715 (5.7)	732 (6.6)	707 (5.3)	716 (5.7)	729 (6.4)	698 (4.8)
Mean MAT Math Score at Entry (Grade Equivalent)	682 (6.2)	691 (6.4)	688 (6.4)	693 (6.5)	680 (6.2)	731 (7.7)	666 (5.8)
Sample Size	371	427	67	101	98	77	84

***The mean of this variable for returnees differs statistically from that of non-returnees at a 0.01 level.

**The mean of this variable for returnees differs statistically from that of non-returnees at a 0.05 level.

*The mean of this variable for returnees differs statistically from that of non-returnees at a 0.10 level.

Table V.7

ADDITIONAL DEMOGRAPHIC AND ECONOMIC CHARACTERISTICS FOR COHORT I RETURNING TREATMENTS

	All Sites		Returnees				
Characteristic	Non-Returnees	Returnees	Boston	Fresno	Portland	San Diego	Seattle
Percent With Limited English Proficiency[a]	7%	11%	0	6%	10%	17%	20%
Percent from Female-Headed Households	52%	50%	68%	48%	53%	47%	40%
Mean Household Size	5.1	5.3	5.2	5.6	5.0	5.7	5.2
Number with Children	3	6	0	0	5	1	0
Percent Ever Worked[b]	43%	33%	33%	41%	43%	28%	17%
Sample Size	371	427	67	101	98	77	84

[a] The percentage of returnees with limited English proficiency differs (at a 0.05 level) from that of non-returnees.

[b] The percentage of returnees who did not have work experience prior to joining STEP differed significantly (at a 0.01 level) from that of non-returnees.

and in Boston, significantly more (68% versus 54%) of the returnees came from female-headed households.

Employment History

Across sites, returnees were much more likely to be those youth with no work experience prior to joining STEP. This is consistent with the notion that youth with few alternative job opportunities were more likely to return to the program. Only 33 percent of the returnees had prior work experience compared to 43 percent of the non-returnees. Significant differences between the returnees and non-returnees were found in the Fresno and Seattle sites on this variable.

School Participation and Educational Expectations Characteristics

Grade of enrollment predicted whether a youth would return for a second summer of STEP. Across all sites and in Fresno, the difference in the grade of enrollment of returnees compared to grade of enrollment of non-returnees is statistically significant. While fewer eighth and 10th graders returned for a second summer, more 9th graders returned for a second summer of STEP. Returnees were also much less likely to have a history of grade retention than non-returnees (25% versus 33%). No significant differences were found between returnees and non-returnees with respect to school absentee rates. Finally, while no differences were found with respect to educational expectations of the two groups across sites, returnees in Portland were much more likely (77% versus 59%) to expect post-high school education than non-returnees in that site.

Data on school participation variables are presented in Table V.8.

Fertility-Related Characteristics

Compared to non-returning treatments, significantly fewer of the returnees across sites reported being sexually experienced prior to joining STEP or, if sexually experienced, significantly fewer reported activity in the two months prior to the start of the STEP demonstration. (This is consistent with the earlier finding that younger participants were more likely to return for a second summer.) This was also true in San Diego and Seattle, where non-returnees were significantly more likely to have had sex prior to joining STEP.

The mean age at which youth became sexually active was roughly equal for both groups of youth, as were the number of males who had been responsible for a pregnancy. Of the 10 treatment females who reported a previous pregnancy, eight returned for a second summer.

Table V.8

SCHOOL PARTICIPATION CHARACTERISTICS AND EDUCATIONAL EXPECTATIONS OF COHORT I RETURNING TREATMENTS

	All Sites		Returnees				
	Non-Returnees	Returnees	Boston	Fresno	Portland	San Diego	Seattle
Grade in School							
8th Grade or Below	57%	50%	85%	33%	45%	52%	48%
9th Grade	34	44	15	60	47	43	45
10th Grade	9	6	0	7	8	5	7
Grade Retention History							
Percent of Full Sample Ever Repeating a Grade[a]	33%	25%	42%	27%	21%	20%	19%
Absenteeism							
Percent Absent 1-10 Days in SY 84-85	69%	73%	76%	67%	74%	78%	76%
Percent Absent 11-20 Days in SY 84-85	21	17	15	20	15	16	18
Percent Absent More Than 20 Days in SY 84-85	10	10	9	13	11	6	6
Educational Expectations							
Percent Planning to Pursue Post-High School Education	71%	73%	54%	71%	77%	80%	80%
Sample Size	371	427	67	101	98	77	84

[a] A significantly lower proportion of the returnees had repeated a grade compared to non-returnees (at a 0.05 level of significance).

In terms of the level of knowledge on birth control issues at baseline, the returnees correctly answered an approximately equal number of questions on general birth control knowledge as the non-returnees. However, returnees had significantly lower scores on the birth control availability questions. No difference in exposure to sex education was found between the two groups across sites. In San Diego, however, returnees were much less likely to have had a previous course in sex education than were non-returnees (73% versus 87%).

Table V.9 presents fertility-related characteristics of the returning treatment sample.

The Returning San Diego Controls

As mentioned earlier, youth who were randomized to the control group in San Diego were given the opportunity to receive work experience in the 1986 summer; 102 did so. The return rate of the controls (85%) was higher than that of the treatments (73%). This pattern is consistent with the decision by many treatments in the five STEP sites to take a full-time job rather than return for a second summer of remediation.

When comparing the baseline characteristics of the returning control youth to those of the non-returning control youth, we found very few differences. Returning control youth had significantly higher educational expectation levels than non-returning controls. The two groups also differed in terms of their general birth control knowledge, with returnees displaying higher levels of knowledge.

Treatment and control returnees in San Diego mirrored each other on basic demographic, educational ability, economic and school participation characteristics. The only significant difference between the two groups was found with respect to the level of knowledge on general birth control issues. On this variable, returning controls significantly outscored the returning treatments. While treatments were less likely to return than controls, the characteristics of the treatment returnees, by and large, mirror those of control returnees.

Tables comparing the characteristics of the returning controls with those of the non-returning controls in San Diego at baseline 1985, as well as tables presenting characteristics of returning controls compared to those of returning treatments are provided in Appendix C.

Table V.9

FERTILITY-RELATED CHARACTERISTICS OF CO

	All Sites		
	Non-Returnees	Returnees	Boston
Percent With Prior Sexual Experience[a]	59%	46%	53%
Percent of the Sexually Active Sample Who Were Sexually Active in the Past Two Months[b]	55%	48%	43%
Percent of Recently Active Sample Never Using Contraception in the Past Two Months[c]	42%	40%	54%
Mean Age at Onset of Sexual Activity for the Sexually Active Sample	11.69	11.64	12.32
Number of Females Ever Pregnant	2	8	0
Number of Males Ever Fathered a Child or Responsible for a Pregnancy	7	5	0
Mean Number Correct on Birth Control Availability Questions (8-item scale)[d]	3.13	2.87	3.10

able V.9

TICS OF COHORT I RETURNING TREATMENTS

	Returnees			
oston	Fresno	Portland	San Diego	Seattle
53%	48%	58%	34%	35%
43%	48%	58%	42%	42%
54%	45%	23%	55%	10%
.32	12.07	11.15	11.41	11.38
0	1	4	1	2
0	1	2	0	2
3.10	2.71	3.01	2.76	2.83

Table V.9 - (cont'd)

FERTILITY-RELATED CHARACTERISTICS OF COHORT I RETURNING TREATMENTS

	All Sites		Returnees				
	Non-Returnees	Returnees	Boston	Fresno	Portland	San Diego	Seattle
Mean Number Correct on General Birth Control Knowledge Questions (12-item scale)	4.86	4.68	5.13	4.74	4.72	4.31	4.52
Percent Believing That Both Partners Should Be Responsible for Birth Control	72%	70%	73%	65%	74%	78%	63%
Percent Who Have Taken a Prior Sex Education Course	64%	61%	44%	55%	74%	73%	53%

[a] The percentage of youth who were sexually experienced prior to joining STEP differed significantly between the returnees and the non-returnees at the 0.01 level.

[b] For youth that were sexually experienced, the percentage of returnees who were sexually active in the two months prior to STEP differed significantly from that of non-returnees (at a 0.01 level).

[c] The percentage of returnees who were non- or irregular users of contraceptives differed from that of non-returnees at a 0.10 level of significance.

[d] The average number of correct responses to the birth control availability questions differed between the returnees and non-returnees at a 0.10 level of significance.

SUMMARY OF PARTICIPANT CHARACTERISTICS

The Cohort II Sample

The youth accepted into the second cohort of STEP are dropout-prone and disadvantaged, based on their academic and economic status. However, in comparison to Cohort I, Cohort II youth are a bit less educationally disadvantaged. Across sites, Cohort II youth scored higher than Cohort I at baseline in both reading and math. Differences in scores for the two groups were significant only in reading, however.

Cohort II is a more diverse group racially and ethnically than Cohort I. Increases in participation by Asians, Hispanics and youth from other racial/ethnic groups account for this difference. There were no differences between the two cohorts with respect to family size and composition or in terms of school participation characteristics.

The Returning Cohort I Treatments

Overall, the treatments who returned for a second summer of remediation and work experience are very similar to non-returnees in terms of their basic demographic, economic, school participation and fertility-related characteristics. There were some exceptions: slightly more of the younger members of the cohort returned for a second summer, and minority youth were more likely to return than were whites and youth from other racial/ethnic groups. Youth with no prior work experience were more likely to return than were youth with prior work experience. Finally, returnees were most likely to be those youth who were not sexually experienced at program entry.

The Returning San Diego Controls

In San Diego, where both treatments and controls could return, we were able to examine whether the treatments who returned differed from the controls who returned. We found the two groups to be virtually identical on most characteristics. Thus, for the most part, the experience of the returning controls probably indicates what would have happened to the treatment returnees in the absence of the program.

VI. FIRST SUMMER IMPACTS--COHORT II AND AGGREGATE SAMPLES

The long-term objectives of STEP are the improvement of high school graduation rates and the reduction of teen parenting rates. STEP strives to achieve these long-term goals by providing work experience, school-year support, remediation and life skills and opportunities classes. The program's short-term success is measured by improvement in academic skill levels and knowledge of contraception and the consequences of adolescent pregnancy.

In this chapter we present analyses of the first summer impacts of STEP with respect to academic and life skills outcomes. As discussed previously, STEP participants entered the demonstration in two waves--the first cohort began the program in 1985 and the results of their first summer of participation were reported earlier (Branch et al., 1986). The second cohort of 14- and 15-year-olds entered STEP in the summer of 1986. The primary focus of the analyses presented here is the impact of STEP participation on Cohort II--youth who entered the program in 1986.[1] In many cases, however, the models are also estimated for the aggregate STEP sample. The aggregate sample consists of both cohorts of STEP youth. Data collected in 1985 for Cohort I are combined with the 1986 Cohort II data. The aggregate results are presented in Appendix C; discussion of these results is limited to differences from the Cohort II results.

As discussed below, Cohort II treatment youth performed significantly better than controls in both math and reading during their first summer of STEP. Participation in STEP did not, however, affect other education-related outcomes (educational expectations, enjoyment of reading and math and re-enrollment in school for the 1986-87 school year). With respect to fertility-related outcomes, treatment youth significantly improved their knowledge of contraception and the consequences of adolescent pregnancy compared to controls. In addition, participation in STEP appears to influence contraceptive use. All of the findings are discussed in detail below.

ACADEMIC PERFORMANCE

In the following sections, the impacts of participation in STEP on a number of measures of academic performance are discussed.

---

[1] An analysis of the determinants of program participation is presented in Appendix A. We found that the best predictor of classroom attendance was the youth's attendance pattern during the preceding school year.

Results presented in this chapter are limited to impacts that can be detected after one summer of program participation.

The primary outcome measures of academic performance are the scaled test scores in reading and math as measured by the Metropolitan Achievement Test (MAT) at the end of the summer. Scaled test scores are the result of the conversion of raw scores (the number of correct answers on the test) to a metric that is more suited to statistical analysis. Although scaled scores are used for analysis, the results are often converted to grade equivalent scores, after the analysis has been completed, to provide an interpretive context for the scaled scores.[2]

We analyze the test scores using two methods. The first method involves a comparison of pre-test to post-test changes in simple unadjusted means for treatment and control youth. These unadjusted means depict actual test scoring without statistical adjustments for any existing pre-treatment differences between treatment and control group youth on important variables. Simple means provide descriptive measures that give the reader a feel for the data, but are not used to determine the program's impact. To determine program impacts, a second analysis method, multiple regression, is used. Regression analysis provides a more precise estimate of the impact because statistical adjustments have been made.[3] For example, a net impact of 10 points in reading indicates that treatment youth would outscore control youth by 10 points at post-test if their baseline characteristics had been identical.

Control Group Losses

For almost every site and every sex or racial/ethnic subgroup, control youth experienced substantial learning losses over the summer. As shown in Table VI.1, Cohort II control youth lost about 20 points (which is equivalent to 1.0 grade) in reading and

---

[2] For the MAT, 10 points on the reading test is equivalent to roughly half a grade level; for math, a change of 20 points represents approximately half a grade level.

[3] The regression coefficient on the treatment variable is an estimate of the impact of STEP controlling for baseline differences between treatment and control youth. The explanatory variables included in the regression equations differ somewhat from the models estimated for Cohort I. Whether the youth lived in a female-headed household and whether the youth's family received public assistance were excluded because: (1) these variables had no effect on the learning analysis outcomes for Cohort I and (2) inclusion of these variables would cause us to lose a substantial number of cases due to missing information.

Table VI.1

MEAN CONTROL GROUP LOSSES IN READING AND MATH IN GRADE EQUIVALENTS BY SITE FOR THE PILOT SAMPLE, COHORT I AND COHORT II

Site	Reading Scores			Math Scores		
	Pilot[a]	Cohort I	Cohort II	Pilot[a]	Cohort I	Cohort II
Overall	na	-0.7	-1.0	na	-0.4	-0.5
Boston	-1.1	-1.4	-0.5	-0.4	-1.0	-0.2
Fresno	na	-0.6	-1.1	na	-0.6	-0.8
Portland	na	-0.2	-1.4	na	0.0	-0.5
San Diego	na	-0.6	-0.5	na	-0.4	+0.1
Seattle	na	-0.7	-1.6	na	-0.1	-1.0
Pinellas County	-0.2	na	na	-0.4	na	na

[a] Only two of the three pilot sites, Boston and Pinellas County, collected data for a randomly assigned control group. Mean test scores and changes were not estimated for the overall sample of control youth in the pilot study.

na = not available

18 points (which is equivalent to 0.5 of a grade) in math.[4] This finding is consistent with earlier research that suggests that summer is a period when learning for many youth stops or decays. In particular, research by Heyns (1978) suggests that this pattern is most prevalent among disadvantaged youth. The control group losses are substantively important because they represent the magnitude of loss that multiply at-risk youth would likely experience. Thus we can assume that these losses would have been experienced by treatment youth in the absence of STEP. The greater the learning losses in the control group, the more effective the program must be to produce gains in learning for the treatment group.

Because the size of the control group loss is substantively important, it was particularly crucial to measure losses free of error and bias. Various unmeasured factors related to the testing situation can affect treatment and control youth differently and artificially inflate the observed control group losses. P/PV took steps to counteract two of these factors in particular--potentially differential motivation to perform well and differential environmental and test proctor effects. To counteract lack of motivation, P/PV developed an incentive scheme designed to motivate both treatment and control youth to do their best at both the pre-test and the post-test. (The details of this incentive were discussed in Chapter III.) Treatment and control youth were tested together in four of the five sites in 1986 to ensure that any biases introduced by test proctors and environmental conditions would influence them equally.[5] We believe these precautions produced comparable testing environments and conditions because, overall, there were no significant differences in pre-test scores between treatments and controls.

Nevertheless, significant control group losses emerged for the overall sample and for most subgroups. Regardless of the precautions taken to reduce the magnitude of control group losses due to testing differentials, losses are large across all years of the demonstration. These results lend support to the hypothesis that the observed losses reflect actual losses that are quite large and worthy of attention.[6]

---

[4] Only in Boston, however, was the magnitude of control groups losses substantially smaller compared to 1985.

[5] In the one site (Fresno) where treatment and control youth were tested separately, the same rooms and proctors were used for testing both groups, and testing occurred at the same time of day.

[6] One final concern, however, was whether these losses represent actual learning loss or are simply an artifact of using the MAT to measure changes in skill levels. P/PV conducted a small study of a group of students who were similar to the STEP

Mean Test Score Changes

Changes in mean test scores from pre-test to post-test for both treatment and control youth in Cohort II are presented in Figures VI.1 through VI.4 and Tables VI.2 and VI.3. Although mean test scores are unbiased estimates of the program's impact, more precise impacts based on multivariate analysis are presented later. Mean reading and math scores at pre-test and post-test were calculated separately for treatment and control groups, by sex, race/ ethnicity and site. We tested whether pre-test to post-test changes within each group were statistically different from zero and whether the amount of change experienced by treatment youth was significantly different from the amount of change experienced by control youth. (Similar analyses were performed for the aggregate sample, the results of which are presented in Appendix Tables C.13 and C.14.)

The results presented in Tables VI.2 and VI.3 indicate that both treatment and control youth in Cohort II experienced statistically significant learning losses in reading over the summer. The control group losses are, however, significantly greater than the treatment group losses.[7] On average, Cohort II treatment youth outperformed control youth by 12.83 points in reading.

In math, treatment youth experienced a significant gain of 7.03 points. This gain, coupled with significant control losses, resulted in an average treatment-control differential of 24.76 points in math.

---

sample and who were participating in another summer program (SWEP) that combined remediation, life skills classes and work experience. These students took the Test of Adult Basic Education (TABE) in addition to the MAT at the beginning and end of their summer program. The implications of these results with respect to summer learning losses are discussed in Appendix B.

[7] For the SWEP study (see footnote 6), P/PV examined the pre-test to post-test changes in scaled scores for both the MAT and the TABE and then converted the results to grade equivalents. The mean score changes for the SWEP sample on the MAT are very close to those for STEP treatment youth over the first summer (in grade equivalents, a loss of 0.4 of a grade in reading and no change in math). On the TABE, however, these students showed gains in both reading and math (equivalent to 0.8 of a grade in reading and 0.3 of a grade in math). This raises the possibility that some of the observed treatment losses might in fact be gains had a different test been used. See Appendix B for further discussion of this issue.

Figure VI.1

MEAN MAT READING SCORES FOR COHORT II AT PRE- AND POST-TEST

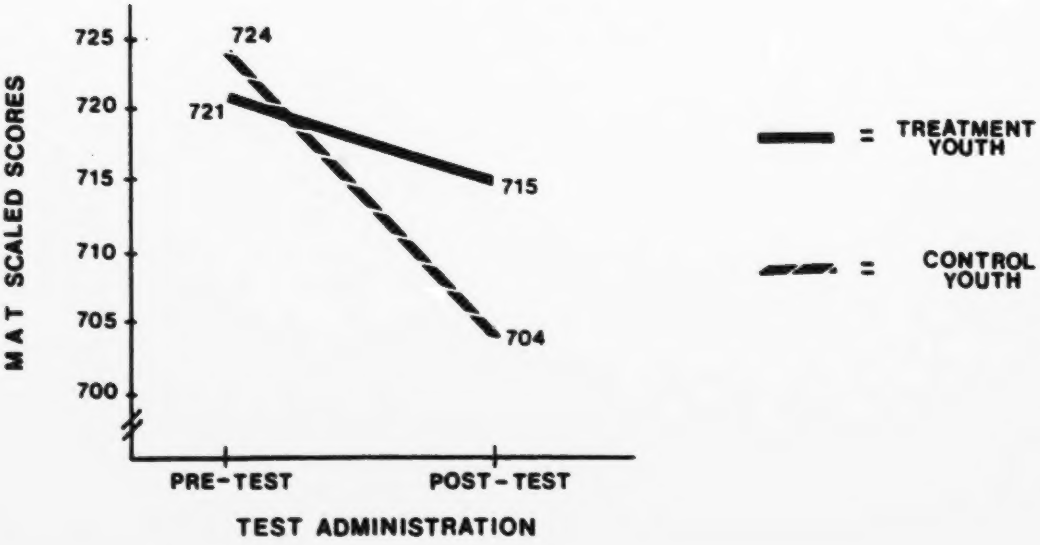

Figure VI.2

MEAN MAT MATH SCORES FOR COHORT II AT PRE- AND POST-TEST

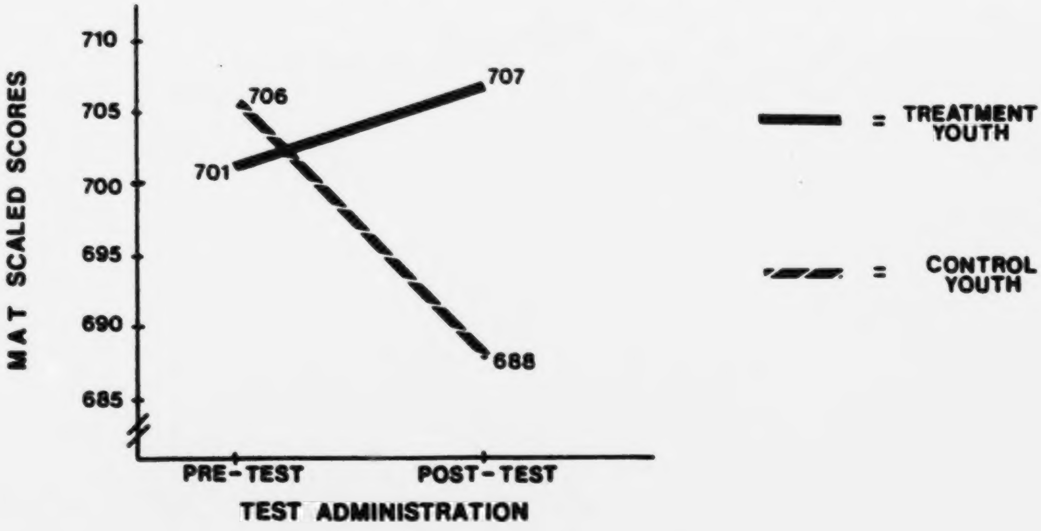

Figure VI.3

PRE-TEST TO POST-TEST CHANGES IN MEAN MAT READING SCORES FOR COHORT II BY SEX AND RACE/ETHNICITY

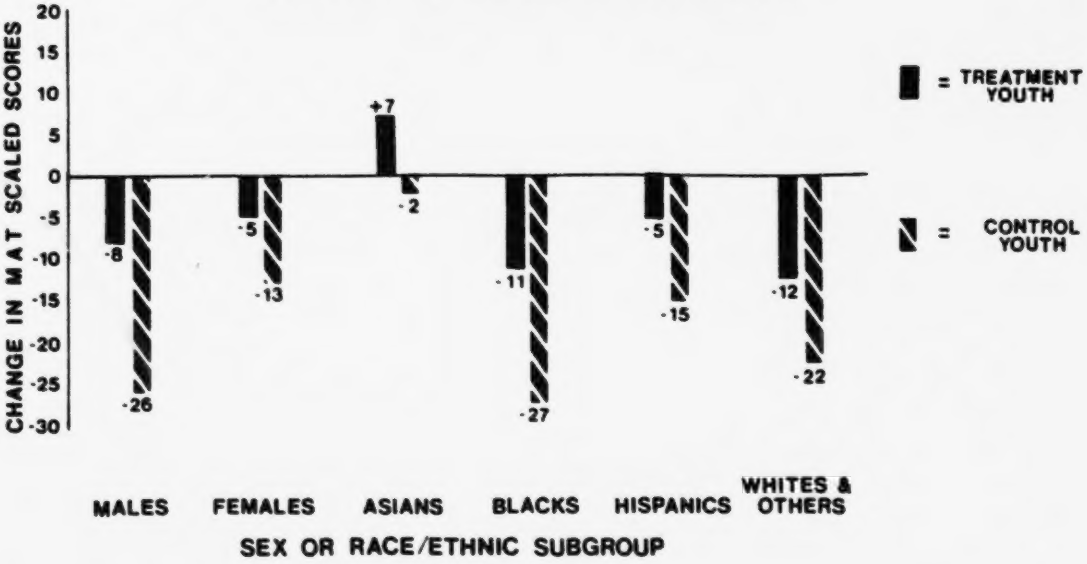

Figure VI.4

PRE-TEST TO POST-TEST CHANGES IN MEAN M A T MATH SCORES FOR COHORT II BY SEX AND RACE/ETHNICITY

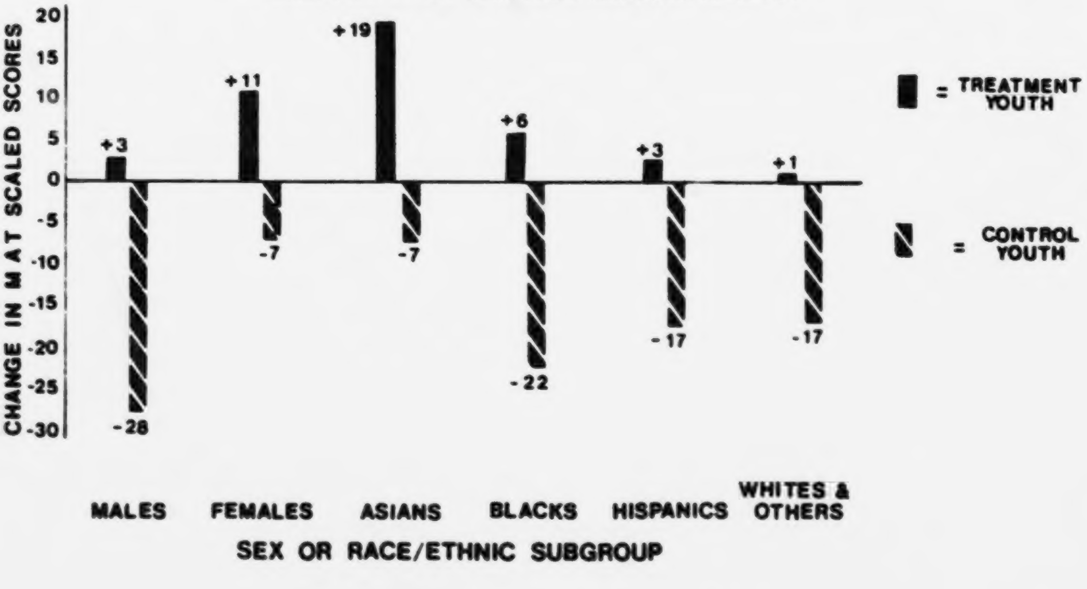

Table VI.2

MEANS OF PRE-TEST TO POST-TEST CHANGES IN READING AND MATH SCORES BY SEX AND RACE/ETHNICITY FOR COHORT II

Group (Sample Size)	Reading Scores: Change from Pre- to Post-test	Reading Scores: Treatment-Control Difference	Math Scores: Change from Pre- to Post-test	Math Scores: Treatment-Control Differences
Total Sample				
Treatment (637)	-6.67***		7.03***	
Control (631)	-19.50***	12.83***	-17.73***	24.76***
Male				
Treatment (297)	-8.25***		2.76	
Control (320)	-26.13***	17.88***	-27.74***	30.50***
Female				
Treatment (340)	-4.90**		10.77***	
Control (311)	-12.68***	7.78**	-7.43**	18.20***
Asian				
Treatment (122)	6.87***		18.88***	
Control (114)	-2.01	8.88**	-7.47	26.35***
Black				
Treatment (307)	-10.52***		5.88*	
Control (301)	-26.73***	16.21***	-22.18***	28.06***
Hispanic				
Treatment (114)	-5.32		2.71	
Control (112)	-15.25***	9.93	-17.15***	19.86***
White & Other				
Treatment (94)	-11.90**		0.68	
Control (104)	-22.29***	10.39	-16.69***	17.37**

***Difference is significantly different from zero at .01 level.
**Difference is significantly different from zero at .05 level.
*Difference is significantly different from zero at .10 level.

Table VI.3

MEANS OF PRE-TEST TO POST-TEST CHANGES IN READING AND MATH SCORES BY SITE FOR COHORT II

Site (n)	Reading Scores: Change from Pre- to Post-test	Reading Scores: Treatment-Control Difference	Math Scores: Change from Pre- to Post-test	Math Scores: Treatment-Control Differences
Total Sample				
Treatment (637)	-6.67***		7.03***	
Control (631)	-19.50***	12.83***	-17.73***	24.76***
Boston				
Treatment (132)	6.56		12.71**	
Control (130)	-9.40**	15.96***	-8.53**	21.24***
Fresno				
Treatment (125)	-16.91***		1.26	
Control (131)	-22.26***	5.35	-25.16***	26.42***
Portland				
Treatment (126)	-12.88***		0.93	
Control (120)	-24.40***	11.52*	-18.72***	19.65***
San Diego				
Treatment (130)	1.58		13.96***	
Control (128)	-9.50***	11.08**	1.72	12.24
Seattle				
Treatment (124)	-11.73**		5.79	
Control (122)	-32.95***	21.22***	-38.99***	44.78***

***Difference is significantly different from zero at 0.01 level.
**Difference is significantly different from zero at 0.05 level.
*Difference is significantly different from zero at 0.10 level.

From the subgroup means in Tables VI.2 and VI.3, we see that the program seems to be more effective for some groups and less for others. Differential program impact for sex, racial/ethnic and site subgroups will be discussed in the context of the multivariate analysis to follow.

Overall Impacts

The regression estimates presented in Table VI.4 confirm that the overall impact of STEP participation was greater for Cohort II than for Cohort I. The net impact of STEP on MAT reading scores for Cohort II was 11.79 points, or 0.6 of a grade equivalent. This is approximately twice the impact that Cohort I treatment youth derived from their first summer of STEP (5.68 points, which is equivalent to 0.3 of a grade). Similarly, the impact of participation in math for Cohort II was nearly double that of Cohort I's first summer results. The net impact on math performance for Cohort II was 23.13, or 0.8 of a grade equivalent. (The net impact for Cohort I had been only 12.64 or 0.3 of a grade equivalent.)

We believe these large impacts on the second cohort are the result of better design and implementation of the remediation component of STEP, and the improvement in the overall program implementation during 1986. (See Chapter IV for a discussion of these improvements.)

STEP had a larger impact on the test scores of Cohort II than it did on those of Cohort I for nearly all sex, racial/ethnic and site subgroups. The only exceptions to this general pattern were: (1) smaller impacts in reading for Cohort II Asians and Hispanics; (2) smaller impacts for Cohort II in Boston in both reading and math; and (3) no difference between Cohort I and Cohort II in the magnitude of impacts in Fresno. As the following discussions of subgroup impacts will show, the effects of STEP were more widespread for Cohort II compared with Cohort I.

Impacts by Sex Subgroups

The results for Cohort II indicate that both males and females derived significant (and statistically equal) benefits from STEP participation in both reading and math. For males, these impacts represent better outcomes than was true for Cohort I males. The stronger math impact for males in 1986 is due to the improved performance of treatment youth over the summer, rather than worse performance of the controls. Control group males in both summers lost about 28 points; treatment group males in the first cohort suffered losses in math, while those in the second cohort had small math gains over the summer. The improvement in reading for males is due to a combination of smaller treatment group losses and larger control group losses than the first cohort experienced. For females, the impact on reading performance remained

Table VI.4

NET IMPACT OF PARTICIPATION IN STEP ON MAT SCALED SCORES IN READING AND MATH FOR COHORT II AND AGGREGATE STEP SAMPLES

Group	Cohort II		Aggregate Sample	
	Reading	Math	Reading	Math
Overall	11.79***	23.13***	9.17***	18.56***
Sex				
Male	15.65***	26.55***	9.21***	20.75***
Female	8.13**	19.88***	9.13***	16.39***
Race/Ethnicity				
Asian	7.63	24.81***	9.93**	20.01***
Black	15.41***	26.46***	7.57***	18.73***
Hispanic	7.96	15.46**	12.89***	19.19***
White & Other	10.11	19.70**	9.91*	15.19**
Site		#	##	###
Boston	11.78**	19.21***	17.27***	29.62***
Fresno	5.32	20.97***	4.64	21.20***
Portland	12.36**	20.54***	2.11	8.53
San Diego	9.59*	13.86*	5.87	7.52
Seattle	20.27***	41.80***	15.43***	25.22***

NOTE: The sample size for the Cohort II analyses is 1,268 and for the aggregate sample analyses. Impacts for each subgroup are computed as the sum of the treatment coefficient plus each subgroup treatment interaction term coefficient. Significance is then determined by the appropriate F-test of the hypothesis that this sum equals zero.

***Indicates that the impact differs statistically from zero at the 0.01 level.
**Indicates that the impact differs statistically from zero at the 0.05 level.
*Indicates that the impact differs statistically from zero at the 0.10 level.

###Indicates that the impact differs with respect to this characteristic at a 0.01 level of significance.
##Indicates that the impact differs with respect to this characteristic at a 0.05 level of significance.
#Indicates that the impact differs with respect to this characteristic at a 0.10 level of significance.

approximately the same as in 1985, while the math impact was somewhat stronger.

The stronger outcomes for males in Cohort II (compared to Cohort I) may be related to the changes that were made in the remediation curriculum. First, the content of the learning modules used in the second summer may have been more interesting and more relevant to males than the first year's curriculum. In addition, males may have responded better to the somewhat more structured, less self-directed approach in Year Two.

Impacts by Racial/Ethnic Subgroups

The biggest improvement among racial/ethnic subgroups from the first cohort to the second occurred for black youth. In 1985, participation in STEP had no impact on the performance of black youth in reading. For Cohort II, the strongest impact on reading scores--15.41 points--was experienced by blacks. In addition, the net impact on math performance more than doubled for blacks in Cohort II compared to their Cohort I counterparts--from about 11 points the first summer to more than 26 points the second summer.

The larger impact for Cohort II blacks in reading is due primarily to larger control group losses in 1986, although treatment youth experienced slightly smaller losses than Cohort I treatments. The math results, however, are due to the combination of treatment group gains of about six points (compared to a two point loss for Cohort I treatments) and somewhat larger control group losses compared to the first cohort (22 points for Cohort II controls versus 17 points for Cohort I controls).

The changes in the remediation curriculum may also explain the better performance of black youth in Cohort II compared to those in Cohort I. The improvements made in curriculum content, structure and computer-assisted instruction may have kept students more on task and engaged in meaningful learning activities.

The strong effects on reading scores for Hispanics and Asians in 1985 are not replicated in the second cohort. In 1985, the strongest impact of participation on reading scores occurred for Hispanic youth, where control group losses were the largest. Asians in the first cohort also experienced a significant benefit from participation with respect to reading. But in the second summer, net impacts in reading for Hispanic and Asian youth were not statistically significant. However, for Asians, although the impact on reading was no longer statistically significant, the decline in the magnitude of the effect was small--equivalent to less than three points.

For Hispanics, the smaller net impact on reading appears to be due primarily to smaller control group losses in Cohort II than experienced by the first cohort. The mean test score changes indicate that Hispanic treatment youth in the two cohorts fared about the same--Cohort I treatment youth lost an average of six points in reading, while Cohort II treatment youth lost an average of five points. Among control youth, however, those in Cohort I lost almost 28 points, while those in Cohort II lost only 15 points.[8] Although the impact on reading was not statistically significant for Hispanics or Asians in Cohort II, the impact for these racial/ethnic groups in the aggregate sample was significant.

The impact of STEP treatment on math scores was uniformly strong across racial/ethnic groups.[9] Asian, black and Hispanic treatment youth in Cohort II all experienced math gains, which represent an improvement in math performance compared to Cohort I treatment youth in the respective racial/ethnic subgroups. At the same time, Cohort II control youth for every racial/ethnic subgroup except Hispanics experienced larger math losses in comparison to Cohort I controls.

In summary, the effects of participation in STEP were more widespread during the second summer, as indicated by the results for Cohort II, compared with the first summer results for Cohort I.

Impacts by Sites

In STEP's second summer, treatment youth in all five sites benefited from participation in STEP. The significant impact of participation in STEP for the first cohort of youth in 1985 was due primarily to the very strong effects in Boston and the somewhat smaller effects in Seattle for reading, and strong effects in Fresno for math. In four of the five sites in 1986 (all except Fresno), the net impact of participation on reading performance was statistically significant. With respect to math, the net impact of STEP participation was significant in all five sites. Comparison of the net impacts within sites from the first to the second cohort, however, can be a bit misleading because of the small sample sizes and factors related to the quality of the program implementation can affect these net measures.

---

[8] This is a puzzling phenomenon that cannot be easily explained.

[9] We cannot reject the hypothesis that these effects are the same for all race/ethnic groups.

Because control group losses represent what would have occurred for treatment youth in the absence of STEP, it is important to use net impacts rather than simple treatment gains or losses as a measure of the program's effect. The net impact on both reading and math for Cohort II was greater than for Cohort I in Portland, San Diego and Seattle. In Boston, however, where Cohort II treatment youth had gains in both reading and math, the net impacts were smaller for Cohort II than for Cohort I, primarily because control group losses were smaller.

Remediation Subsites Differences

Remediation occurred at three universities in Boston and five high schools in San Diego. In Seattle, both morning and afternoon classroom sessions were held. In the first summer there were very large and significant differences in the pre-test to post-test changes across the five San Diego remediation subsites. Although subsite differences were detected in Seattle and Boston in the first summer as well, these differences were much smaller than those observed in San Diego and, with the exception of reading change scores in Boston, were not statistically significant.

Given the subsite differences observed for the first cohort, we examined whether any also occurred for Cohort II. At each of the three sites where remediation occurred at multiple sites (or morning vs. afternoon sessions), we calculated the pre-test to post-test changes in mean reading and math scores for each of the remediation subsites. Mean scores for all remediation subsites are presented in Table VI.5.

There are no significant subsite differences in the mean test score changes. However, the sample sizes within each subsite are so small that only very large differences are likely to be statistically detected.[10] Thus, the pattern of test score changes across subsites may still be an indication of uneven implementation of the remediation component in these sites.

None of the subsite estimates are statistically different from one another (which is not unexpected given the small sample sizes for each subsite), and the numbers themselves are not too different. In San Diego, in particular, these results provide evidence of a more even implementation of STEP during 1986 than during 1985. Mean score changes for San Diego subsites in the first summer ranged from losses of about 57 points (in both

---

[10] We also estimated net impacts for remediation subsites using regression techniques. The results of this analysis indicate that within each site, the impacts across the remediation subsites are statistically equal.

Table VI.5

MEANS OF PRE-TEST TO POST-TEST CHANGES IN READING AND MATH SCORES BY REMEDIATION SUBSITE IN BOSTON, SAN DIEGO AND SEATTLE FOR COHORT II

	Reading Scores			Math Scores		
Site	Pre-test	Post-test	Change	Pre-test	Post-test	Change
Boston						
UMass	696.00	706.62	10.62	661.04	680.27	19.23**
Northeastern	683.30	694.34	11.04	667.09	672.43	5.34
BU	714.25	710.98	-3.27	687.35	701.40	14.05*
San Diego						
Hoover	743.89	744.93	1.04	770.70	793.04	22.33**
Kearny	708.96	718.42	9.46*	745.50	760.13	14.63
Lincoln	731.88	723.68	-8.20	722.32	717.00	-5.32
Morse	717.63	718.59	0.96	700.04	716.07	16.03
San Diego	742.56	747.33	4.77	735.67	756.41	20.74
Seattle						
Morning	721.04	713.86	-7.18	691.71	699.21	7.50
Afternoon	732.21	716.38	-15.83*	700.94	700.90	-0.04

NOTE: None of the pre-/post-test changes differed significantly (at a 0.10 level) across subsites or time of day.

***Indicates that the impact differs statistically from zero at the 0.01 level.
**Indicates that the impact differs statistically from zero at the 0.05 level.
*Indicates that the impact differs statistically from zero at the 0.10 level.

reading and math) to gains of about 14 points in reading and 37 points in math. As indicated in Table VI.5, the range of average score changes for Cohort II in San Diego was substantially narrower, from an 8 point loss to a 9 point gain in reading and from a 5 point loss to a 22 point gain in math. It is clear from these results that the implementation of remediation in San Diego was more successful in the second summer.

There is no clear pattern, however, of subsite differences in Boston. Students attending the University of Massachusetts subsite appear to have benefited the most. Students attending classes at Northeastern University did well in reading, but not as well in math, while the Boston University students did well in math, but had small losses in reading.

In Seattle, youth who attended classes in the morning had small gains in math (as opposed to no change for those in afternoon remediation) and reading losses that were half the size of those in afternoon remediation. This suggests, but does not confirm, that afternoon classes may be less effective than morning remediation.

Aggregate Sample Outcomes

The net impacts of STEP participation, presented in Table VI.4, were significant for the overall aggregate sample, for males and females and virtually all race/ethnic groups. There were, however, significant differences in the impact of participation across sites. In the aggregate sample, participation significantly affected only post-test reading scores in Boston and Seattle and math scores in Boston, Fresno and Seattle. Since the aggregate sample is comprised of the combined cohorts, these results represent an average of the relatively modest Cohort I results discussed in the report on the 1985 experience and the stronger Cohort II results discussed above.

Impact of Teacher Characteristics

Classroom observations during the first summer of STEP indicated some possible variation in the quality of instruction being delivered. Although observational data suggested that characteristics of remediation teachers might influence the impact of STEP participation, no systematically collected data that could be used to address this issue were available. In 1986, however, we collected information from remediation teachers that can be used to investigate whether characteristics of instructors affect the outcome measures.

Under the assumption that STEP participants would be more responsive to teachers from similar backgrounds, the STEP sites attempted to recruit a cadre of teachers with a diverse racial/ethnic mix. In addition, it was expected that effective STEP instructors would be teachers with previous experience relevant to the situation they would face in STEP and teachers who viewed their instructional style as compatible with STEP's instructional approach. The questionnaire completed by remediation teachers was designed to measure prior experience and compatibility with STEP's instructional approach as well as to collect basic demographic information.

Relevant teacher characteristic variables were added to the basic learning equation models in order to assess their relative impact on learning outcomes. Specifically, four variables were added to the basic model: (1) whether the sex of the student and teacher matched; (2) whether the race/ethnicity of the student and teacher matched; (3) a composite measure of teacher experience; and (4) the compatibility of STEP philosophy with teaching style. The experience measure was created from each teacher's response to three questions: whether they had previous experience with a special needs population; whether they had previous experience with individualized, skills-based instruction; and whether they had previous experience with computer-assisted instruction. The compatibility between STEP philosophy and teaching style was measured with four response categories ranging from "completely" to "not at all." In situations where youth had separate reading and math teachers, the characteristics of the appropriate teacher were included in the model.

The results of this analysis indicate that none of these teacher variables had a significant impact on either reading or math post-test scores. This means that the teacher characteristics we measured appeared not to affect the learning outcomes. Although observational data suggest that there were variations in the quality of instruction received by STEP youth, the effect of these variations, if any, are more subtle than we were able to measure. The issue of quality of instruction and teacher characteristics will be discussed further with respect to life skills outcomes.

Hours Analysis

In all of the analyses discussed to this point the impact of STEP was determined by comparing outcomes for treatment youth with those for control youth. Estimating the impact of participation in this way treats all STEP participants as if they were exposed to the same amount of remediation and life skills. Although Cohort II treatment youth attended an average of 100 classroom hours, the previous analyses treated youth who attended relatively few classes in the same manner as those with perfect attendance. To determine whether increased exposure to remediation resulted in better outcomes, we estimated a model that provides an estimate of the incremental impact of each hour of treatment on post-test scores.[11]

The results of this analysis indicate that, on average, each additional treatment hour results in an increase of 0.12 points on the MAT reading test and an increase of 0.23 points on the MAT math test (both of these estimates are statistically significant at the 0.01 level). Although this may seem small, it translates into 1.2 points in reading and 2.3 points in math for every additional 10 hours of remediation. Participation in 100 hours of STEP (as was the average for Cohort II treatments) results in an impact of approximately 12 points in reading and 23 points in math. These impacts translate to approximately 0.6 of a grade equivalent in reading and 0.8 of a grade equivalent in math.

Attitudes Toward Education

One might hypothesize that a program that increases an individual's skills in reading and math might also influence his/her attitudes toward school, particularly toward reading and math. Additionally, if the youth perceives that s/he can be successful in school, his/her expectations for educational attainment may increase. Alternatively, attitudes toward school may decline over the summer given that youth have been engaged in reading and math activities during that time without a break from the regular school year. Measures of STEP participants' enjoyment of reading and math and their expectations for higher education were collected at baseline and endline to investigate these hypotheses.

Students were asked how much they enjoyed reading and math with five response categories ranging from "very much" (coded as 1) to "not at all" (coded as 5). (Thus a higher number connotes less enjoyment.) For both treatment and control youth in Cohort II, the mean level of enjoyment of reading at baseline was approximately 2.7, while the mean for enjoyment of math was about 2.5. These mean levels did not change significantly over the summer for either group, and there was no treatment-control difference in the amount of change that occurred.

---

[11] Since the same unmeasurable variables that affect how well an individual performs at post-test are also likely to predict how many hours of classes s/he attended, we estimated the impact of treatment hours on test scores using two-stage least squares techniques. The same set of exogenous variables included in the previously discussed net impact equations were used to predict treatment hours and post-test scores simultaneously. The estimated model included all Cohort II youth who are in the learning analysis sample, with the number of treatment hours for control youth set to zero.

Educational expectations were measured by asking youth the highest grade or year in college they expected to complete. To examine whether these expectations changed as a result of STEP, we compared the percentage of youth who expected any post-high school education at baseline with the corresponding percentage at the end of the summer. At baseline, 71 percent of treatment youth and 74 percent of control youth indicated that they expected to continue their education after high school. These percentages increased slightly for both groups over the summer--up 3 percentage points for treatment youth and 4 percentage points for control youth. These increases are not statistically significant, nor is the increase for treatments significantly different from the increase for control youth.

We also estimated multivariate models for the impact of STEP on enjoyment of reading and math and educational expectations. The results of the regression models indicate that participation in STEP had no impact on enjoyment of reading or math.

Two models for examining the impact of STEP on educational expectations were estimated. The results of both analyses indicate that participation in STEP had no impact on educational expectations. The first model provided an estimate of the probability that a youth expects post-high school education. The second model regresses an ordinal measure of educational expectations[12] at the end of the summer on treatment status, baseline educational expectations and the set of explanatory variables included in previous learning analysis equations.

Return to School Information

The final educational outcome measure examined was re-enrollment in the 1986-87 school year. The number of youth in the STEP sample attending school at the beginning of the school year following STEP was obtained from school district records. The percentages of Cohort II treatment and control youth who returned to school in the fall of 1986 are presented in Table VI.6.

In general, the percentage of youth who returned to school was high--more than 96 percent of both treatment and control youth. That such a high percentage returned to school is not surprising given that virtually all of these youth are younger than 16, the legal age to leave school.

---

[12] This measure of educational expectations ranges from less than high school graduation (which is coded as 1) to four years of graduate work (coded as 10). Thus an expectation to graduate from high school is coded as 2, one year of college as 3 and so forth through four years of graduate work.

Table VI.6

COHORT II YOUTH WHO RE-ENROLLED IN SCHOOL, FALL 1986, OVERALL AND BY SEX, RACE/ETHNICITY AND SITE

Characteristic	Treatment	Control	Difference
Overall	96.5%	96.9%	-0.4%
Sex			
Male	95.8%	96.8%	-1.0%
Female	97.1	97.1	0.0
Race/Ethnicity			
Asian	96.4%	97.9%	-1.5%
Black	96.1	97.8	-1.7
Hispanic	98.0	91.5	6.5***
White & Other	96.1	99.2	-3.1*
Site			
Boston	97.0%	98.8%	-1.8%
Fresno	94.9	93.0	1.9
Portland	92.9	98.0	-5.1**
San Diego	99.4	96.3	3.1*
Seattle	98.6	98.6	0.0

***Difference between treatments and controls is significantly different from zero at 0.01 level.

**Difference between treatments and controls is significantly different from zero at 0.05 level.

*Difference between treatments and controls is significantly different from zero at 0.10 level.

To determine whether participation in STEP affected re-enrollment, we tested whether re-enrollment rates differed by treatment status. For the overall sample of Cohort II youth, no significant difference in the percentage who returned to school was found using either logit analysis or difference-in-means analysis.

In addition, for most subgroups there was no difference by treatment status in the percentage of youth who returned to school. The exceptions are: (1) among Hispanic youth and youth in San Diego, treatment youth were significantly more likely to return to school, and (2) among the category "whites and others" and youth in Portland, control youth were more likely to return to school. Although these few subgroups showed significant differences between treatments and controls, the actual percentages are still very high for all groups. Regardless of which subgroup is considered, more than 90 percent of Cohort II youth returned to school following the first summer of STEP.

Summary

In summary, the impact of STEP on learning, as measured by post-test MAT scores, was positive and strong for Cohort II. The net impacts of participation for Cohort II were 11.79 points in reading and 23.13 points in math. These impacts are equivalent to 0.6 of a grade in reading and 0.8 of a grade in math.

These learning impacts for Cohort II are stronger and more widespread than for Cohort I's first summer of STEP. In math, Cohort II treatment youth of both sexes, all race/ethnic groups and all five demonstration sites benefited from participation in STEP.

In reading, Cohort II males and females, black youth and youth in four of the five sites who participated in the STEP treatment benefited significantly. The impacts for Hispanics and Asians, and for youth in Fresno, however, were not significant.

None of the teacher characteristics that were measured had a significant influence on either reading or math post-test scores. If characteristics of the remediation instructors affected learning, these measures were not able to capture the differences.

The other education-related outcomes measures examined--enjoyment of reading and math, educational expectations and returning to school--were not affected, either positively or negatively, by treatment status. Mean levels of enjoyment of reading and math did not change over the summer among either treatment or control youth. The slight increase in the percentage of youth expecting post-high school education was not significant.

LIFE SKILLS OUTCOMES (LSO)

Improvement in the level of basic skills addresses one of the major factors related to dropping out of high school. The life skills component of STEP addresses the other major factor in this process--adolescent parenthood. The focus of the life skills component is the relationship between decisions (social and sexual) youth make at 14 or 15 and future outcomes in the areas of education, work and family. In the long run, STEP aims to reduce teenage pregnancy and parenting rates. Impacts on these outcomes, however, are difficult to detect by the end of the first summer because there is not a long enough follow-up period. The short-term impacts that can be detected over one summer of STEP relate to changes in knowledge of and attitudes toward contraception and the consequences of teenage parenting, the level of sexual activity and the use of contraceptives. The sections that follow present the findings of analyses of the first summer fertility-related outcomes for Cohort II and the aggregate sample of STEP youth.

The life skills analysis sample, used for estimating impacts on knowledge, included 1,302 youth who have non-missing values on the fertility-related variables.[13] For the Cohort II life skills analysis sample, the sex, age and racial/ethnicity compositions are not significantly different for treatments and controls. To investigate the possibility of differential attrition by treatment group status, a logit model was estimated. The results indicate that treatment-control group status does not affect the probability of being in the life skills analysis sample. Thus, we are more confident that the impacts represent unbiased and consistent estimates of the treatment effect on the life skills analysis sample.[14]

---

[13] The variables included in the life skills multivariate analysis models are: treatment status, sex, age, race/ethnicity, and baseline measures of family size, birth order, English language problems, absenteeism, knowledge of contraception, knowledge of availability of contraception, knowledge of the consequences of adolescent pregnancy, responsibility for birth control and demonstration site dummy variables. In addition, dummy variables for non-response on measures of female-headed household, ever had sex (at baseline) and ever had a sex education course were included along with dummy variables for these three measures in order to limit sample attrition.

[14] A similar attrition analysis was completed for the aggregate sample with similar results.

Contraceptive Knowledge

In order to assist youth in making responsible social and sexual decisions, a portion of the life skills curriculum concentrates on the presentation of factual information related to avoidance of pregnancy. A visit to a comprehensive health facility is included as part of the curriculum to heighten awareness of the accessibility of health care, including birth control.

We hypothesized that participation in the life skills classes would result in increased knowledge about contraception over the course of the summer program. To test this hypothesis, we developed two composite measures of birth control knowledge. The first measure consists of the number of correct responses to 12 true or false statements about various methods of contraception. An examination of the mean number of correct responses to these statements (presented in Appendix Table C.15) indicates the need these youth have for birth control information. At baseline, both treatment and control youth in Cohort II were able to correctly answer an average of only five of the 12 statements.

The impact of STEP on knowledge of birth control was estimated using a multiple regression model, the results of which are presented in Table VI.7. These results indicate that treatment youth significantly increased their contraceptive knowledge by an average of 2.64 questions more than control youth. This represents an improvement of approximately 50 percent over the mean baseline level.

The second indicator of the impact of STEP on contraceptive knowledge consists of eight questions on the availability of various birth control methods. For each question, youth were asked to indicate whether the method is available without seeing a doctor. On average, Cohort II youth were able to correctly answer only three of these eight questions at baseline.

The results of the multivariate analysis, presented in Table VI.7, indicate that STEP significantly increased youths' knowledge of the availability of birth control. At the end of the summer, treatment youth correctly answered an additional 1.74 questions more than controls. Given that the mean level at baseline was 3, this outcome also represents a 50 percent increase in knowledge.

The significant impact of STEP on knowledge of contraceptive methods and knowledge of contraception availability observed in the overall sample holds for both sexes, for all racial/ethnic groups and across all sites. The magnitude of the impacts are similar for males and females, and for black, Hispanic and white youth, but are quite a bit larger for Asian youth. The stronger impact for Asian youth may be partially due to the relatively low level of contraceptive knowledge possessed by Asian youth at

Table VI.7

NET IMPACT OF PARTICIPATION IN STEP ON BIRTH CONTROL KNOWLEDGE VARIABLES FOR COHORT II AND AGGREGATE BY SEX, RACE/ETHNICITY AND SITE

	Cohort II			Aggregate Sample	
Group	Contraceptive Knowledge[a]	Knowledge of the Availability of Contraception[b]	Knowledge of the Consequences of Teen Pregnancy[c]	Contraceptive Knowledge[a]	Knowledge of the Availability of Contraception[b]
Overall	2.64***	1.74***	0.25***	2.51***	1.72***
Sex		##			#
Male	2.60***	1.94***	0.29***	2.42***	1.78***
Female	2.63***	1.51***	0.22***	2.61***	1.64***
Race/Ethnicity	###	###		###	###
Asian	3.89***	3.13***	0.32***	3.16***	2.62***
Black	2.31***	1.39***	0.30***	2.18***	1.48***
Hispanic	2.32***	1.54***	0.19**	2.78***	1.83***
White & Other	2.46***	1.53***	0.13	2.62***	1.44***
Site	###	###	###	###	###
Boston	1.24***	0.63***	0.14**	1.54***	0.84***
Fresno	2.51***	1.42***	0.30***	3.06***	2.01***
Portland	2.75***	1.95***	0.16**	2.34***	1.56***
San Diego	4.75***	3.31***	0.47***	3.83***	2.87***
Seattle	1.90***	1.42***	0.26***	1.75***	1.28***

NOTE: This table presents selected results from separate regressions for each of the dependent variables for the overall sample and each of the subgroups presented. Lack of comparable data on knowledge of the consequences of teen pregnancy precluded estimation of impacts for the aggregate sample.

[a] Mean number of correct responses on a 12-item scale.
[b] Mean number of correct responses on an 8-item scale.
[c] Mean number of correct responses on a 2-item scale.

***Indicates that the impact differs statistically from zero at the 0.01 level.
**Indicates that the impact differs statistically from zero at the 0.05 level.
*Indicates that the impact differs statistically from zero at the 0.10 level.

###Indicates that the impact differs with respect to this characteristic at a 0.01 level of significance.
##Indicates that the impact differs with respect to this characteristic at a 0.05 level of significance.
#Indicates that the impact differs with respect to this characteristic at a 0.10 level of significance.

baseline. Although the average STEP youth could correctly answer five of the 12 knowledge questions and three of the eight availability questions at baseline, Asian youth in STEP correctly answered only three of the 12 knowledge questions and two of the eight availability questions at baseline. Thus Asian youth had more room to improve their knowledge than did other racial/ethnic groups in the STEP sample.

These impacts were significantly different across sites. The impacts on Cohort II in Fresno, Portland and Seattle are very similar to each other. However, the impacts in San Diego are about twice as large, while those in Boston are only about half the size of the coefficients estimated for the other three sites.[15] Classroom observations suggest that the divergent results for San Diego and Boston may reflect differences in the delivery of LSO instruction in these sites, where teachers' previous experience handling this kind of sensitive material varied dramatically.

For example, the two life skills instructors for Cohort II in San Diego taught the life skills component of STEP during the first summer and taught primarily sex and drug education during the school year. Based on classroom observations, these San Diego instructors were outstanding. On the other hand, none of the seven Cohort II LSO teachers in Boston taught life skills during the first summer of STEP or taught health or sex education during the school year.[16] Only three of the seven were currently teachers by profession. The experience of the instructors was more mixed in the other sites.[17] This variation in the characteristics of the instructors may have contributed to the site-by-site variation in impact.

---

[15] The range of the net impacts on each of the measures of contraceptive knowledge was narrower for Cohort I than that observed for Cohort II. On the knowledge of contraceptive methods measure, the Cohort I range was 2.19 (from 1.71 in Boston to 3.90 in Fresno) compared to a range of 3.51 for Cohort II (from a low of 1.24 in Boston to a high of 4.75 in San Diego). For the impact on the knowledge of birth control availability measure, the between site range was 1.74 for the first cohort (the low, in Boston, was 0.88 while the high, in Fresno, was 2.62). The comparable range for the second cohort was 2.68; Boston again had the smallest impact, 0.63 and the largest occurred in San Diego, 3.31.

[16] Although the regular LSO instructors were not sex education teachers, two of the more technical sessions of the curriculum that dealt with anatomy and reproduction were delivered by specialists from a local family planning program.

[17] In Seattle, the two Cohort II LSO instructors who returned for a second summer of STEP were teachers by profession, while the third instructor was a counselor new to STEP. None of the

To further investigate the possibility that instructor characteristics affected the life skills outcomes, measures of teacher characteristics were included in the appropriate regression equations. In particular, the race/ethnicity and sex of the instructors, whether they are health/sex education teachers and whether they had taught LSO during the first summer of STEP were included.

The regression analysis indicated, however, that none of these characteristics affected the impact of STEP on knowledge of contraception or knowledge of the availability of contraception. Although observational data suggested variation in the delivery of the LSO curriculum, the teacher characteristics that were subject to measurement may not capture the critical differences that were observed across sites.

The scheduling of LSO classes in Boston may have contributed to the smaller impacts observed in that site. Unlike the other four sites, where twice-weekly LSO classes were incorporated into the daily remediation schedule, Boston youth attended LSO classes in the afternoons (twice a week) after already spending three hours in remediation classes during those mornings.

Analyses of the impact of STEP on the two contraceptive knowledge measures were also completed for the aggregate sample of STEP youth; the results are also presented in Table VI.7. Again the impacts are significant overall and for each of the subgroups. As was true for the learning analyses discussed above, the aggregate results represent an average of the impacts observed for each of the two cohorts. Since the Cohort II results were generally a little stronger than those for Cohort I, the aggregate coefficients are slightly smaller than those reported for Cohort II.

One final indicator of the impact of STEP on fertility-related knowledge was the participants' understanding of the consequences of adolescent pregnancy. Cohort II youth were asked two questions--whether girls who became mothers were likely to continue their education or drop out of school, and the average cost of

---

Cohort II instructors in Portland and Fresno had previous STEP experience. Both of the Portland instructors were teachers, one of them in health education. Three of the four Fresno instructors were teachers, but in areas other than health or sex education. Instructors indicated only their current profession on the questionnaires, thus whether any teachers had previous experience teaching health or sex education was not measured.

pregnancy and the first year of a baby's life. At baseline, the average Cohort II youth correctly answered less than one of these two questions. Treatment youth significantly increased their understanding of these consequences over the summer, while control youth knowledge did not change. The net impact of STEP on this fertility-related measure is reported in Table VI.7. For Cohort II youth overall, the estimated impact is 0.25. With the exception of the category "whites and others," the impact of STEP on this outcome measure was significant for all the subgroups for which impacts were estimated.

Sexual Behavior

The sensitivity of the questions on sexual behavior led to greater non-response than on other questions. Of the 1,516 Cohort II youth who attended the pre-testing session and completed the baseline questionnaire, 7 percent failed to answer the question on their previous sexual experience. Only those youth with previous sexual experience were asked the remaining questions that dealt with sexual behavior and contraceptive use. Thus the sample available for analyzing sexual behavior and contraceptive use is smaller than that used in the previous analyses. The sample size is further depleted by non-response on these items at the end of the summer.

Of the Cohort II youth who answered the sexual experience question at both measurement points, 43 percent of treatment youth and 45 percent of control youth indicated at baseline that they were sexually experienced (Table VI.8). Among Cohort II youth who had ever had sex at baseline, approximately 54 percent (Table VI.8) indicated at baseline that they were recently sexually active, that is, that they had sex within the previous two months. Over the summer, in the absence of the program, both the percentage of youth reporting that they are sexually experienced and the percentage reporting recent sexual activity increased--by 6.8 percentage points and 5.6 percentage points, respectively (as seen among control youth). This increase is not surprising given the effects of aging and the amount of free time these youth have during the summer.

Given the short duration of STEP and the timing of the sessions that deal with sexual behavior, we do not expect to find an impact on sexual behavior during the first summer. Multivariate analyses of these measures indicate, however, that STEP may have influenced sexual behavior. We estimate that participants were significantly less likely--by 24 percent--to report being

Table VI.8

BASELINE TO ENDLINE CHANGES IN THE PERCENTAGE OF COHORT II YOUTH WITH SEXUAL EXPERIENCE, HAVING SEX IN THE PREVIOUS TWO MONTHS AND USING CONTRACEPTIVES AT LAST INTERCOURSE

	Sexual Experience[a]		Sexually Active[b]		Contraceptive Use[c]	
	Baseline	Change Over Summer	Baseline	Change Over Summer	Baseline	Change Over Summer
Treatments	42.5%	5.7%***	53.7%	10.6%***	37.8%	15.2%***
Controls	45.1	6.8 ***	54.4	5.6 *	33.7	6.9 **
Difference		-1.1		5.0		8.3
Males						
Treatment	60.8%	4.6%***	54.6%	9.0%***	41.2%	17.6%***
Control	62.2	8.4***	57.0	6.7 **	32.4	9.7 **
Difference		-3.8 *		2.3		7.8
Females						
Treatment	26.4%	6.8%***	51.9%	13.6%**	31.2%	10.4%
Control	26.8	5.0 ***	48.1	2.6	36.8	0.0
Difference		1.7		11.0		10.4

[a] Percentage of youth reporting that they have ever had sex.
[b] Percentage of sexually experienced youth reporting that they have had sex in the previous two months.
[c] Percentage of sexually experienced youth reporting use of contraception the last time they had sex.

***Indicates that the impact differs statistically from zero at the 0.01 level.
**Indicates that the impact differs statistically from zero at the 0.05 level.
*Indicates that the impact differs statistically from zero at the 0.10 level.

sexually experienced at the end of the summer than were controls holding constant whether they had ever had sex at baseline.[18]

However, the impact estimate on the probability of having had sex in the last two months is essentially zero (treatment youth were 14% more likely to report recent sexual activity) and is not significant.[19]

The results of these two analyses are inconsistent. We are not sure what effect, if any, STEP had on sexual activity over the summer because we get conflicting results using two different measures. We had not expected, however, STEP to affect sexual activity during the program given the concurrence of the program delivery and the measurement of the behavior. Thus, whether participation in STEP affects sexual behavior cannot be determined until follow-up interviews are completed.

The final fertility-related measure of the impact of STEP participation for Cohort II youth is change in contraceptive use behavior. As with the other sexual behavior measures, only a small sample is available for analysis. Of the Cohort II youth who are sexually experienced, response rates on the contraceptive use questions were high. At baseline, 97 percent of the sexually experienced youth responded to the question of whether they had used birth control the last time they had sex. The response rate for this question was somewhat lower at the end of the summer--91 percent of the sexually experienced youth answered the contraceptive use question.

As shown in Table VI.8, among Cohort II youth who responded to the contraceptive use question, 38 percent of treatments and 34 percent of controls indicated at baseline that they had used some form of birth control the last time they had sex. At the end of the summer, while 41 percent of controls reported using contraception during their last intercourse, the increase in reported contraceptive use among treatments was even greater--8.3 percentage points more than for control youth.[20]

---

[18] Some have speculated that this finding could be due to measurement error. It is generally believed that boys tend to over-report sexual activity. Exposure to the treatment might result in boys, who (over)reported they were sexually active at baseline, reporting that they had never been sexually active at endline. An investigation of this possibility suggests that the treatment effect reported here is not due to this type of measurement bias.

[19] Both impact estimates were derived using logit analyses. Both analyses were conducted on youth who answered the question "Have you ever had sex?" at endline.

Logit analysis of the probability of using contraception during last intercourse indicated a significant impact on the probability of using contraception among those youth who were sexually active during the program period. Treatments are one and a half times more likely to have used contraception during their last intercourse than controls. Thus although the evidence is inconclusive regarding the impact of STEP on sexual activity, STEP significantly affected contraceptive use among youth who were sexually active during the summer.

Although the samples are small, we also examined contraceptive use for sex and site subgroups. There was no difference in estimated probability of contraceptive use for males and females. However, we did find a significant difference across sites. In San Diego and Boston, sexually active treatment youth were significantly more likely than controls to have used contraception during their last intercourse--more than three and a half times more likely in San Diego and almost twice as likely in Boston. In the other sites there was no significant impact on contraceptive use.[21]

Finally, analyses of sexual behavior and contraceptive use for the aggregate sample were not conducted because the presentation of the behavioral measures in the questionnaires differed across the two cohorts. The changes made to the questionnaire from the first to the second summer of STEP affected the response rates to these questions. Thus the behavioral information we have is not comparable across cohorts and precludes estimating impacts for the aggregate sample of youth.

---

[20] Although 623 youth answered the contraceptive use question at baseline, and 642 answered the question at endline, only 491 Cohort II youth provided information at both times -- 230 treatment youth and 261 control youth. Among these treatment youth, the percentage who indicated that they had used birth control during their last intercourse increased by 15.2 percentage points over the summer (from 37.8% at baseline to 53.0%), which is statistically significant at the 0.01 level. The percentage of control youth who indicated contraceptive use also increased significantly (at the 0.05 level) over the summer, but only by 6.9 percentage points.

[21] Although this looks unusually large, the marginals indicate that while the percentage of control youth using contraception in these two sites did not change over the summer, the percentage of treatment youth reporting contraceptive use increased from 35 to 54 percent in Boston and from 25 to 78 percent in San Diego. Further, we reviewed the sample sizes for these two sites used in the logit equation that was run for all sites. While they are not large, they are sufficient (112 in Boston and 89 in San Diego).

Summary

Participation in STEP clearly influenced the fertility-related knowledge of these youth. With respect to knowledge of contraceptive methods, treatment youth correctly answered an average of 2.74 questions more than controls at the end of the summer. Similarly, treatment youth increased their knowledge of the availability of contraception by an average score of 1.74 more than control youth. Both of these represent increases of approximately 50 percent over the mean baseline scores. Understanding the consequences of adolescent pregnancy also increased among treatment youth in Cohort II.

Treatment youth--males and females from all racial/ethnic groups and from all five sites--significantly improved their knowledge compared to control youth on the three measures of fertility-related knowledge. These results for Cohort II are even stronger than those for Cohort I in 1985. Finally, none of the measured characteristics of LSO instructors affected the outcome measures of fertility-related knowledge.

From our analysis of the first summer for Cohort II, we are not sure what effect, if any, STEP may have had on sexual behavior over the summer because we get conflicting results using two different measures of sexual activity. The major short-term impact of STEP appears to be a substantial increase in the use of contraception among Cohort II youth who were sexually active during the summer.

VII. SCHOOL YEAR AND SECOND SUMMER IMPACTS FOR COHORT I

Cohort I youth have now completed an academic year and many have returned for their second summer of STEP instruction and work experience. Thus, we can examine what effect the first summer of STEP had on the behavior of treatment and control youth during the intervening school year. We can also examine who returned in the second summer of STEP and how they performed. Because controls are not offered a second summer of work experience, information on them was not collected during the second summer (except in San Diego).[1] Thus, we primarily describe how the reading, math and life skills of treatment youth changed over this second summer and for the full 15 months.

SCHOOL DATA ANALYSIS

STEP was designed to improve students' school performance by providing them with needed remediation, increasing their understanding of the consequences of not finishing school, giving them the knowledge and ability to prevent pregnancy and, perhaps, helping them to have a positive school experience. To measure STEP's short-term impact on school performance, we examine five outcome measures for the academic year 1985-86: attendance, test scores, credits earned, grade promotion and dropout behavior. One school outcome that we do not consider is grade point average, which is too difficult to interpret without information on the content and difficulty of courses taken. Without adjusting grade point average for such factors, perverse results are quite possible. If, for instance, STEP treatment youth took college preparatory courses while controls did not, it might appear that STEP had a detrimental effect on grade point averages. Because of these potential problems, grade point average is not examined.

Overall, we found that the first summer of STEP had modest effects on the behavior of Cohort I during the school year 1985-86. Treatment youth appeared to be somewhat more likely to be promoted to the next grade and to score slightly higher in reading on standardized test scores than were control youth. Other outcome measures were unaffected by STEP. Before we present these results in greater detail, however, we briefly discuss issues that lead one to interpret them cautiously.

---

[1] Data will be collected for controls during subsequent school years and with the follow-up interviews.

Discussion of Data Limitations

No one measure provides a complete picture of school performance, and no one measure is without limitations. Thus, we investigate the effect of STEP participation on all five variables noted above and look for a confluence of impact results, so that we can be more confident in our conclusions. Below is a brief summary of the issues encountered when analyzing each of the variables.

Attendance

Data on absenteeism are notoriously inaccurate. For example, work done in the Boston school district indicates that the recorded number of days present and the recorded number of days absent do not always add up the total number of days in the school year. In our sample of Boston students, the sum adds to the total number of school days in only 46 percent of the cases. Eighty percent are within plus or minus 10 days of the true total.

If measurement errors are large, the effect of the program on attendance also needs to be large to achieve statistical significance. Problems also occur if the accuracy of the data varies systematically among the sites.

Differences in attendance-taking practices among sites cause us to suspect that the quality of data does differ by site. Portland took attendance in each of eight periods during the day; the other sites only in homeroom.[2] The method of recording attendance also varied. In Boston, teachers recorded a student's presence or absence, but elsewhere, only absences were recorded. Thus, the accuracy of the attendance data is likely to differ systematically among the sites. Although differences in attendance-taking practices led us to interpret cross-site findings cautiously, we did take steps to minimize the problem. In particular, we adjusted for systematic errors across sites by analyzing attendance in a regression setting where specific site influences can be controlled statistically.[3] We also checked the robustness of the pooled site results by examining absenteeism separately by site.

---

[2] To make Portland attendance data comparable to that of the other sites, we used attendance during the first period of the day.

[3] Another problem occurs if the degree of random measurement error varies significantly among the sites. The presence of measurement error makes an estimated impact appear to be closer to zero than it actually is. If the degree of the measurement error varied among the sites, the site impacts would be biased by different amounts. Thus, differences in the site impacts would not only indicate differences in the STEP impact by site, they would also reflect the differences in measurement error. There is no way of ascertaining decisively whether the

Standardized Test Scores

Within a given school district, analysis of test score data was straightforward. We encountered complications primarily when we attempted cross-site analysis. Sites administered different standardized tests, and the scores from these tests were not directly comparable. Direct analysis of scaled scores had to be confined to a site-by-site analysis. The sample within a site, however, was relatively small (approximately 150 treatments and 150 controls), making it difficult to observe statistically significant effects.

To conduct a cross-site analysis of test scores, we had to convert the scores to "normal curve equivalents (NCEs)." NCEs are similar to test score percentiles, ranging from 0 to 100, but unlike percentiles, they are a linear measure that can be subjected to regression analysis. The transformation of the scaled scores to normal curve equivalents enabled us to pool test data across sites and analyze the test data of the entire sample. The larger sample size increases our ability to detect program impacts, but translation may add additional unexplained variation to the data, which may somewhat inhibit our ability to observe significant program impacts.

To verify the accuracy of the impacts based on the pooled analysis of test-score normal curve equivalents, we also conducted site-by-site analyses of scaled scores. The site-by-site examination of test scores is relatively uncomplicated analytically. However, two of our sites--San Diego and Portland -- presented us with specific problems. In San Diego, students are tested only every other year. Thus, only sample members who were in the seventh, 9th or 11th grades were tested. Though this selection process does not bias the results presented in this report, the number of sample members for which there is data was reduced by more than half.[4]

---

degree of measurement error varied by site, however, we found that the variance in absenteeism was similar among all sites except Seattle where the variance was 82 percent smaller than that of the other sites.

[4] When only a portion of the sample is included in analysis, one must take care that the estimated impacts are not biased. If the reason an individual is included is related to the treatment, results can be biased. In San Diego, STEP participation did not affect whether we obtained a test score for the school year immediately after the first summer of STEP. A student's grade for that year was determined prior to STEP enrollment. In other

An alternative way of determining whether the effect on reading and math skills persisted after an academic year--possible only in San Diego--is to compare the treatment and control scores on the tests that Cohort I youth took at the beginning of their second summer of STEP.[5] The advantage of using these test scores is that a greater percentage of the sample took these tests than took the in-school standardized test (80% versus 43%). However, this approach has two disadvantages. First, since participants took the STEP test at the end of June, and the in-school tests in April, the STEP MAT test results are less comparable to scores on in-school tests from other sites. Second, because only some of the sample members returned for a second summer of STEP, the sample of returnees may not represent the full cohort, or there might have been a differential pattern of returning for treatment and control group members. Such differential attrition could bias our impact estimates. Thus, we examined the difference in the test scores for only the returning treatments and controls to investigate the robustness of our in-school test results.

In Portland, not all the sample members took a standardized test. Portland students are tested annually between grades three and eight. However, between 9th and 12th grade, only students who have not yet met a minimum level of competence in reading or math are tested. This procedure provided us with test scores for approximately 40 percent of the Portland sample members. Thus in Portland, we could address the question: how did STEP affect the reading and math skills of those who have not already reached the minimum level of competency (or who are in the eighth grade or less --4 % of Portland's sample). However, we could not determine the impact on test scores for the average Portland sample member, as we could in the other sites.[6] Thus, when we aggre-

---

words, the selection criterion--being in grade 7, 9 or 11--is independent of STEP. Thus, no bias occurs in analyzing the data from the first school year. However, a similar type of test score analysis conducted on the data next year would potentially suffer from sample selection bias because STEP may affect grade progression.

[5] This option is not available in the other sites because controls did not return and were not tested in these other sites.

[6] This sample selection will not bias our analysis this year because whether an individual has reached the minimum level of competence as of the school year after their first summer of STEP is independent of STEP. However, the level of the youth's math and reading skills is probably not independent of his/her STEP participation in any year after the first. Thus, next year we must correct for the selection process statistically or reframe our question to: given that a student had not met the minimum required level of competency as of the intervening school year,

gated the test scores across sites (using normal curve equivalent), we had to omit Portland.

Credits

As a third measure of STEP's impact on school-year progress, we collected data on the number of high school credits earned by each student. Because students who were in eighth grade or below did not earn high school credits, they were excluded from the analysis of credit accumulation. This decreases the sample size for the credit analysis by 8 percent, slightly increasing the difficulty in finding statistically significant impacts.

Grade Progression

The fourth measure of school-year performance is grade progression, which is influenced by a school district's philosophy of grade promotion. Promotion in one school district may not signify the same academic advancement as promotion in another. Therefore, grade promotion measures a student's progress through a school system, not necessarily his/her skills acquisition or academic achievement.[7]

The more lenient a district's promotional philosophy, the smaller the estimated STEP impact is likely to be. At the extreme, suppose all students were promoted no matter what academic level they attained by the end of the year. There would be no difference in the proportion of treatments and controls who passed; therefore, we would estimate that STEP had no impact on grade progression. Such an outcome, however, would not necessarily imply that STEP did not affect academic achievement.

Thus, we must be especially cautious when using grade progression to compare impacts across sites. We attempted to minimize this problem by using regression analysis and by analyzing promotion separately by site to check the pooled results. Nevertheless, we must base our assessment of STEP's effect on educational progress on analyses of both accumulated credits and grade promotion.

---

how many years does it take to achieve the competency level for STEP treatments versus controls?

[7] The proportion of students who are promoted differs somewhat in our five sites. Among the controls, the percentage who passed are the following: 65.0 percent in Boston, 68.9 percent in Fresno, 80.1 percent in Portland, 92.6 percent in San Diego and 76.4 percent in Seattle.

Dropout Behavior

Accurate statistics on dropouts are very difficult to find. Schools often do not know if students have dropped out because they have not or could not have (if under 16 years old) formally withdrawn. Thus, most school districts remove students from the active enrollment list based on irregular attendance. However, the manner in which school districts record such withdrawals varies considerably. In addition, since schools qualify for federal funds based on the number of children they serve, the quality of dropout and enrollment data must be suspect.

In an attempt to examine dropout behavior while recognizing these problems, we used two definitions of dropping out. Under the first definition, an individual was deemed to have dropped out if the school so reported or if P/PV so discovered.[8] Withdrawal because of involuntary expulsion or institutionalization was not counted as dropping out, but voluntary withdrawal, such as enlistment in the armed forces, was. These measures are likely to underestimate the number of dropouts for two reasons. First, the school system may have inaccurately reported that the student was still enrolled. Second, we were unable to reach some of the youth whose records looked suspicious. Some of them may have been dropouts.

In the second definition, we included everyone who was a dropout according to the first definition plus individuals for whom there was no apparent or identifiable reason for withdrawal. This measure of dropping out probably overstates the number who dropped out because some of these individuals could have transferred to other schools without informing their previous schools. Thus, neither measure accurately gauges the size of the problem, but we attempted to establish boundaries for the true measure with the two we examined. However, if both measures reflect STEP's impact--measured as the difference in dropout behavior between treatments and controls--potential inaccuracies in the dropout rate may not be as large a problem as they first appear.

In the next section, we present the findings from our analyses. The impact that we observed this year, however, may be smaller than the impact will be in future years and for future cohorts for two principal reasons. First, Cohort I youth still had to complete the program by participating in their second summer of STEP. Thus, the results presented in the following section are mid-program results. Second, as mentioned in the last chapter, we found that the effect of STEP on Cohort I at the end of their first summer was rather small relative to the effect on Cohort II. This may result from improvements in the program after its first year of operation. For example, school-year support activities were minimal for Cohort I; P/PV strengthened and focused school-year support for Cohort II. With these caveats in mind, we present the findings.

---

[8] P/PV attempted to contact any student for whom school data were missing. Primary and secondary contacts were called to locate youth. When enrollment in new schools was reported P/PV requested data from these schools. Through these contacts P/PV sometimes learned that a student had dropped out.

Results

In this section, we examine the impact of STEP on each variable --attendance, reading and math test scores, credits earned, grade promotion, and dropout behavior--for Cohort I youth during the school year immediately following their first summer of participation. The discussion focuses primarily on estimates derived from regression analyses that pooled data from all the sites. The results of the site-by-site analyses conducted on these outcomes (presented in Appendix C) are examined in order to determine the robustness of the pooled findings. We find that both the pooled and site-by-site analyses show that the small impact STEP had on the first cohort at the end of its first summer did not greatly alter school-year behavior or performance the following year. STEP did, however, increase the probability of a youth passing a grade and being promoted.

Attendance

The number of days in the school year varies slightly among STEP school districts. We standardized for these differences by dividing days absent by the total number of days in the school year for each student to derive the percentage of the school year that the student missed.

Table VII.1 shows that for the sample as a whole, STEP had a numerically small and statistically insignificant effect on absenteeism.[9] In the absence of the program, Cohort I youth would have been absent 13.81 percent of the school year, on average. STEP enrollment reduced the annual absenteeism rate by only 0.91 percentage points, which represents a 6.6 percent reduction.

The impact of STEP on absenteeism did not vary greatly by sex, race/ethnicity or site. The effect was greatest for Hispanics (a 14.0% reduction, or -2.63 percentage points). This is consistent with the finding last year that STEP was most effective for Hispanics. However, we cannot statistically reject the hypo-

---

[9] This analysis does not include sample members from San Diego, where attendance is not kept on a student's historical file.

Table VII.1

THE ESTIMATED IMPACTS OF STEP ON ABSENTEEISM AND TEST SCORES IN READING AND MATH DURING THE INTERVENING SCHOOL YEAR BY SEX, RACE/ETHNICITY AND SITE FOR COHORT I

Group	Percentage of the Year Absent[a]	Reading Scores in NCEs[b]	Math Scores in NCEs[b]
Overall	-0.74%	1.44*	1.53
Sex			
Male	-0.70	0.70	1.15
Female	-0.79	2.22*	1.99
Race/Ethnicity			
Asian	1.01	4.03*	1.94
Black	-0.57	1.55	1.51
Hispanic	-2.63	-2.26	-0.16
White & Other	-0.46	4.40	5.11
Sites			
Boston	-2.79*	1.30	0.59
Fresno	1.00	0.04	2.50
Portland	-1.47	na	na
San Diego	na	1.32	2.60
Seattle	0.27	3.29**	0.79
Sample Size	928	675	673

NOTE: Of the 1,576 members of Cohort I, approximately 20 to 25 percent had missing data on one or more of the variables used in the regressions. In specific regressions, the sample was further reduced because the outcome measure was missing in a site or for particular groups of youth.

[a] San Diego did not provide us with absenteeism data.

[b] In Portland, test data were available only for those who had not already achieved a minimum level of competency. Thus, the test score regressions exclude Portland because their test score data were not comparable to that of the other sites. In addition, only 40 percent of the youth in San Diego had test data because tests are only administered there in seventh, 9th and 11th grades. Tests of differential sample attrition by treatment status indicated, however, that attrition rates were equal for treatments and controls. We also tested to see if the impact statistically varied by sex, race/ethnicity and site. None of these subgroup impacts differed from each other at even a 0.10 level.

The control group means were: 13.81% for percentage of the year absent, 39.80 for the reading score and 48.25 for the math score.

***Indicates that the impact is statistically different from zero at the 0.01 level.

**Indicates that the impact is statistically different from zero at the 0.05 level.

*Indicates that the impact is statistically different from zero at the 0.10 level.

thesis that the impact was the same across racial groups. Similarly, while the difference in the absenteeism rates between treatments and controls was greatest and statistically significant in Boston (-2.79 percentage points or a 20.2% reduction), we cannot reject the hypothesis that the impact was equal across sites.

Test Scores

The third and fourth columns of Table VII.1 show that, after the first school year, the modest increases in reading and math scores seen at the end of the first summer of STEP (as reported in Branch et al., 1986) did not persist.[10] STEP appeared to increase standardized test scores (measured in normal curve equivalents that range from 0 to 100) only slightly. In the absence of the program, the average scores would have been 39.80 in reading and 48.25 in math. STEP participants scored only 1.44 higher in reading and 1.53 higher in math than they would have scored in the absence of the program.

Because only 40 percent of the Cohort I youth in San Diego were given standardized tests, we tested the robustness of these results by substituting the scores that returning treatments and controls in San Diego achieved on their 1986 STEP-administered pre-test for school-administered test scores. By doing so, we obtained data on 80 percent of the San Diego Cohort I youth. The estimated impact using this sample was even smaller, and both the reading and math impacts were not statistically significant. (These impacts are presented in Appendix Table C.20.)[11]

While Table VII.1 indicates that impact on reading scores varied statistically by race/ethnicity, this effect is not robust. The effect disappeared when the analysis was redone using the alternative San Diego data.

To further investigate the impact of STEP on reading and math, we conducted a site-by-site analysis of scaled scores. As discussed earlier, scaled scores from different standardized tests are not comparable and cannot be pooled across sites. The results of this analysis are presented in Table VII.2. In order to compare the estimated impacts in this table across sites, we divided the

---

[10] Because Portland's test data were not comparable to that of the other sites, Portland is excluded from this analysis.

[11] To get some idea of whether the test scores of returnees differed systematically from non-returnees, we compared their school-administered (CTBS) test scores. We found no statistical difference between the two groups.

Table VII.2

THE ESTIMATED IMPACTS OF STEP ON SCALED TEST SCORES DURING THE INTERVENING SCHOOL YEAR BY SITE FOR COHORT I

Site (Standarized) Test Use)	Scaled Reading Scores		Scaled Math Scores	
	Estimated Impact as a Percentage of the Control Group Mean	Sample Size	Estimated Impact as a Percentage of the Control Group Mean	Sample Size
Boston (MAT)	0.64	198	0.37	195
Fresno (CAT)	0.03	203	0.53	204
San Diego (CTBS)	1.19	98[a]	1.24	102[a]
(MAT)	-0.14	187	-0.13	187
Seattle (CAT)	1.90	175	0.58	172
Portland (PALT)	0.27	101[b]	0.38	118[b]

[a] In San Diego, the school district administered the CTBS to only 7th, 9th and 11th graders. Thus, only 40 percent of our sample took the CTBS. Several weeks later, returning treatments and controls were given the MAT before their participation in the second summer of STEP. We estimated the impact of STEP using both sets of data.

[b] In Portland, only individuals who had not already reached a certain minimum level of competency (40% of our Portland sample) were given a test. Thus, the sample in Portland is not comparable to the samples in other sites.

***Indicates that the impact is statistically different from zero at the 0.01 level.
**Indicates that the impact is statistically different from zero at the 0.05 level.
*Indicates that the impact is statistically different from zero at the 0.10 level.

estimated impact by the mean scaled score of the control group in each site. Thus, for example, treatments in Boston scored 0.64 percentage points higher in reading than controls, whereas in Fresno they scored only 0.03 percentage points higher. All of the site-by-site impacts were quite small and none were significant.

Portland's test data were not comparable to the data received for other sites because Portland administered its tests only to those students who had not reached a minimum level of competency in reading and math. Thus, we did not pool Portland youth with the other youth in our overall analysis of test scores. The estimated effect of STEP on Portland youth is found only in Table VII.2. As in the other sites, STEP did not have a significant effect on test scores. Treatment youth who had not reached the minimum level of competency were no more likely to do so after their first summer of STEP participation than were control youth.

Overall, it appears that the first summer experience with STEP did not lead the treatment youth of Cohort I to score significantly higher on their standardized reading or math tests during the school year than they would have otherwise.

Credit Accumulation

The total number of credits needed to graduate varies considerably among school districts. To compare academic progression across sites, we standardized the reported number of credits to a common metric. In order to examine the proportion of credits earned toward graduation, we divided the number of credits earned by the total number of credits needed to graduate.[12] As men-

---

[12] Not only did the metric by which credits were reported differ by site, but whether the site reported annual or cumulative credits also varied. Fresno and Portland reported annual credits, whereas the other three sites reported cumulative credits. However, because the number of cumulative credits at the beginning of the 1985-86 school year were the same for treatments and controls, the difference in cumulative credits at the end of the year between treatment and controls is an estimate of STEP's impact, just as the difference in annual credits between treatments and controls is an estimate of STEP's impact. To estimate this impact based on a pooled sample, the following specification was estimated:

$$\text{cumulative credits} = a + bX + cT$$

$$\text{annual credits} = d + eX + cT$$

where X is a vector of personal characteristics, and a through e are regression coefficients to be estimated. The estimate on the treatment variable, T, is constrained to be the same in the two equations.

tioned earlier, because eighth graders do not earn high school credits, the analysis here is restricted to students who were in the 9th grade or higher during the 1985-86 school year.

Table VII.3 indicates that STEP participation had no effect on credit accumulation. Treatments and controls earned practically the same percentage of credits. Treatment youth who were in the 9th grade or higher earned only 0.9 percent (0.17 percentage points) more credits toward graduation than did control youth. We also do not find any significant variation in the impact by sex, race/ethnicity or site. The site-by-site estimates, similarly, show small insignificant results.

Probability of Promotion

STEP did appear to significantly increase the chance that treatment youth would pass their grade. Analysis of the probability of promotion indicates that an average STEP participant would be 22.1 percent (5.27 percentage points) less likely to fail than he or she would have otherwise been. This constitutes a substantial increase since approximately 24 percent of Cohort I youth would have failed in the absence of the program.

The largest impact was experienced by Hispanics. We estimate that STEP enrollment significantly increased the probability of promotion for Hispanics by 39.9 percent (or 12.13 percentage points). The large impact for Hispanics is consistent with last year's finding that the program appeared to be more effective for them. Although the impact for Hispanics appears to be larger than average, we cannot reject the hypothesis that it is the same as the impact for other racial/ethnic groups.

Probability of Dropping Out

As discussed earlier, it is very difficult to determine who is a dropout. We considered two definitions of dropouts: a strict definition that included only students who were recorded as dropouts by the school district; and a less restrictive definition that included all recorded dropouts plus individuals who had withdrawn for an unstated reason. Some of these additional people are individuals with whom the schools and P/PV have simply lost touch, but some are probably true dropouts.

Unfortunately, the difference in what these two definitions measure is unclear, since much of the variation among sites is caused purely by the way they record withdrawals. Portland, for example, records almost all withdrawals as due to "other reasons" --not due to transfer. Thus, in Portland almost no one is explicitly a dropout. In San Diego, on the other hand, unless the school knows that a student has transferred or was specifically exempted from the truancy laws, the youth is recorded as "Dropped (unknown)." Thus, dropout rates varied considerably by

Table VII.3

THE ESTIMATED IMPACTS OF STEP ON CREDITS EARNED AND GRADE PROMOTION DURING THE INTERVENING SCHOOL YEAR BY SEX, RACE/ETHNICITY AND SITE FOR COHORT I

Characteristic	Percentage of Graduation Credits Earned	Probability of Promotion[a]
Overall	0.17%	5.27%
Sex		
Male	-0.62	2.73
Female	0.98	5.81
Race/Ethnicity		
Asian	0.58	14.48
Black	-0.41	8.80
Hispanic	1.38	16.06*
White & Other	1.02	-5.21
Sites		
Boston	0.74	2.09
Fresno	-0.82	3.53
Portland	-1.24	9.39
San Diego	0.08	3.55
Seattle	2.38	-3.47
Sample Size	1,003	1,168

NOTE: Of the 1,576 members of Cohort I, approximately 20 to 25 percent had missing data on one or more of the variables used in the regressions. In specific regressions, the sample was further reduced because the outcome measure was missing for particular youth. We also tested to see if the impact statistically varied by sex, race/ethnicity and site. None of these subgroup impacts differed from each other at even a 0.10 level.

The mean percentage of graduation credits earned for the control group is 18.71% and the mean probability of promotion for the control group is 76.14%.

[a] The probability of promotion was estimated using logit analysis.

***Indicates that the impact is statistically different from zero at the 0.01 level.

**Indicates that the impact is statistically different from zero at the 0.05 level.

*Indicates that the impact is statistically different from zero at the 0.10 level.

site and definition. According to the strict definition, the dropout rates were 4.9 percent in Boston, 2.9 percent in Fresno, 0.7 percent in Portland, 15.0 percent in San Diego and 5.5 percent in Seattle. However, according to the less restrictive definition, the rates were 4.9 percent in Boston, 32.7 percent in Fresno, 23.5 percent in Portland, 15.0 percent in San Diego and 6.2 percent in Seattle.

When we examined the dropout rates of treatments and controls, no significant differences were found using either definition. Thus, STEP enrollment did not appear to affect the probability that an individual would drop out of school. Because we admittedly have no good method of measuring dropout rates, questions about school retention may be better addressed by examining grade promotion and graduation, both of which are well defined.

Summary of Analysis of School Data

We examined the effect that one summer's participation in STEP had on the subsequent year's school behavior and performance of Cohort I youth. In particular, we examined attendance, standardized test scores, credits earned, grade promotion and dropout behavior. Overall, we find that the only school outcome affected by STEP was grade promotion. Treatments were 22.1 percent less likely to fail than were controls. It appears that the program had an especially large impact on promotion rates for Hispanics; however, statistically, the impact does not differ by racial/ethnic group. Although the impact on promotion is encouraging, we do not correspondingly find that treatments earned more credits. Thus, we are cautious about drawing strong conclusions from this one positive result.

The lack of strong effects on school outcomes is not entirely surprising given that the impact of STEP was small for Cohort I at the end of its first summer. At that time, treatments scored only a quarter of a grade equivalent greater in math and reading than controls. Perhaps stronger impacts (such as we see for Cohort II), a second summer of STEP participation, and strong school-year support are required to produce substantial effects on school outcomes. We will be able to investigate these possibilities after the 1986-87 school year is completed.

CHARACTERISTICS OF COHORT I RETURNEES

All Cohort I treatment youth were invited to participate in a second summer of work experience, remediation and life skills instruction in 1986. Only the San Diego site decided to offer control youth a second summer job. As discussed in Chapter III, because of the extreme difficulties encountered when attempting to re-test control youth from the pilot during the summer following their participation, P/PV decided not to re-test control youth who were not being offered second summer jobs.

The returning sample looks very similar to the original Cohort I treatment group. The one exception is that the younger members of Cohort I were somewhat more likely to return for a second summer of STEP. Although youth who were 14 in the first summer were more likely to return, the baseline ability level of the returnees was nearly identical to the original sample. Mean math scores at baseline for the two groups were identical; in reading, the returning youth averaged only 0.1 of a grade equivalent lower than the original group.

In San Diego, the returning control youth also resembled the original control group with one notable exception. The returning control youth appear to be somewhat less academically proficient than the original group, scoring 0.2 of a grade equivalent lower in reading and 0.5 of a grade equivalent lower in math on the baseline MAT. Although these mean scores for returning controls are lower than those for the original sample, they closely resemble the mean scores for returning treatment youth in San Diego. (For a more complete discussion of the Cohort I returnees vis-a-vis the original Cohort I sample, see Chapter V.)

Table VII.4 presents selected information describing the Cohort I returnees at the beginning of their second summer in STEP. As the 1986 pre-test MAT scores indicate, Cohort I treatment youth were performing at about the 6.4 grade level in reading and the 7.2 grade level in math. With close to half of these youth enrolled in the 10th grade or higher during the 1985-86 school year, they are still performing well below their grade level of enrollment.

Although the Cohort I returnees look very similar to the original treatment youth with respect to their program entry characteristics, some notable changes in the characteristics of these youth occurred during the year subsequent to the first summer of STEP. When they began STEP, 73 percent of Cohort I returnees reported that they expected to obtain some post-high school education. At the beginning of the second summer, however, only 68.6 percent expected post-high school education. In addition, problems with school attendance appear to have increased during the intervening school year. While 27 percent of these youth had reported more than 10 absences during the 1984-85 school year, more than 38 percent reported an excess of 10 absences during 1985-86.[13] From the school data analysis presented above, we know that the attendance behavior of treatments and controls did not differ.

---

[13] The school data indicate that 61 percent of the returnees were absent more than 10 days in the 1985-86 school year. Thus it appears that self-reported absenteeism is severely under-reported.

Table VII.4

SELECTED CHARACTERISTICS OF RETURNING COHORT I TREATMENT YOUTH AND SAN DIEGO CONTROL YOUTH AT BEGINNING OF SECOND SUMMER, BY SITE

	All Treatments	Treatments					San Diego Controls
		Boston	Fresno	Portland	San Diego	Seattle	
Mean MAT Reading Score	727.65	726.42	719.68	733.99	737.33	721.96	742.75
Grade Equivalent	6.4	6.3	6.0	6.7	6.9	6.1	7.2
Mean MAT Math Score	713.52	699.70	713.52	704.03	740.99	710.44	743.48
Grade Equivalent	7.2	6.7	7.2	6.8	8.1	7.0	8.3
Grade During SY 1985-86							
9th or Below	51.3%	88.1%	35.6%	42.9%	55.8%	46.4%	49.0%
10th or Above	48.7	11.9	64.4	57.1	44.2	53.6	51.0
Percent Absent More Than 10 Days During SY 1985-86	38.2%	35.0%	52.0%	37.0%	35.9%	26.8%	28.3%
Percent Expecting Post-High School Education	68.6%	44.8%	65.3%	75.5%	77.9%	75.0%	69.6%
Knowledge of Contraception	6.72	6.55	7.35	7.19	6.31	5.91	5.42
Knowledge of Availability of Contraception	4.52	4.52	4.83	4.48	4.51	4.23	4.00
Percent with Sexual Experience	51.1%	56.7%	59.4%	58.2%	45.5%	33.3%	38.2%
Percent Ever Pregnant	6.1%	10.4%	5.0%	9.2%	1.3%	4.8%	-
Percent Ever Fathered a Child	3.7%	1.5%	9.9%	1.5%	2.6%	2.4%	-
Sample Size	427	67	101	98	77	84	102

The attendance of the treatment youth was no better or worse than the controls' attendance during the 1985-86 school year. Finally, the percentage of treatment youth with sexual experience increased by about 6 percent (from 46% to 52%) during the year between the beginning of STEP and the beginning of the second summer.

DESCRIPTION OF SECOND SUMMER AND 15-MONTH OUTCOMES

All Cohort I youth who returned for a second summer of STEP took the MAT Intermediate Level reading and math tests and completed questionnaires at the beginning and end of their second summer. These data provide the primary indicators of second summer outcomes for returning STEP participants. In addition to the analysis of second summer outcomes, outcomes over the 15-month period are examined.

The analyses are limited, however, by the lack of second summer data on the Cohort I control group. Thus we primarily describe the changes in outcome measures over the second summer and 15-month period. The changes over the second summer are calculated as the difference between the 1986 pre-test and post-test outcomes levels. The changes over the 15-month period are calculated as the difference between the 1985 pre-test and the 1986 post-test levels. We test to determine if the estimated changes are equal to zero. This approach is used to examine both learning and life-skills related outcomes.[14]

The primary exception to this strategy is the analysis of San Diego data. Because the San Diego site decided to offer SYETP jobs to control youth during the second summer, questionnaire and test data were collected for these returning San Diego control youth. Thus in San Diego, we are able to estimate net impacts of program participation on the various outcome measures.[15] Although we cannot generalize the San Diego results to the demonstration as a whole, these results provide some evidence as to the effectiveness of the program.

---

[14] For the learning analysis, a multivariate analysis using two-stage least squares techniques provided estimates of the marginal impact of an additional hour of treatment (i.e., remediation and life skills classes) on post-test reading and math scores.

[15] Multivariate regression models that parallel those estimated for first summer impacts were estimated for San Diego returnees.

Description of Learning Outcomes

Second Summer Outcomes

Mean MAT reading and math scores for each of the four test administrations are presented in Table VII.5 for the overall group of returning treatment youth and for returning San Diego treatment and control youth. Although the returning youth are only a subset of the original Cohort I sample, the first summer changes for the returnees mirror those experienced for the original cohort (-5.46 vs. -8.55 in reading and 1.34 vs. -0.24 in math).

The differences in second summer scores indicate that treatment youth also experienced a significant loss in reading (equivalent to 0.9 of a grade) during the second summer. In math, however, the loss was small, equivalent to 0.1 of a grade, and not statistically significant. These outcomes represent significantly larger reading losses than the same youth experienced during their first summer of STEP, and significantly larger reading and math losses than the reading loss and math gain experienced by Cohort II youth during the 1986 summer. Without comparable data on a control group, though, it is difficult to assess STEP's effect. We can, however, determine whether increased exposure to STEP results in better post-test scores. The results of this analysis indicate that the marginal impacts of an extra hour of classroom training are a statistically insignificant 0.35 points in reading and a statistically significant 1.06 points in math in the second summer.[16]

The control group data collected in San Diego provide an alternative way of gauging STEP's effect. Over the course of the STEP demonstration we have consistently found that control youth suffer substantial learning losses during the summer, and that treatments consistently outperform control youth, even though

---

[16] Two-stage least squares techniques were used to estimate these impacts. The parallel analysis reported for Cohort II's first summer outcomes included control youth (with treatment hours set to zero for controls) in the estimated model. Theoretically, the model that includes treatments and controls and the one that includes only treatment youth are the same. When we conducted the treatment-only analysis for Cohort II we obtained larger estimates of the impact than when control youth were included in the model, although we could not reject the hypothesis that the estimates obtained from each model were equal. The treatment-only model may not fully control for the relationships between motivation, participation and test scores and thus the marginal impact of each second summer treatment hour, presented here, may be overestimated.

Table VII.5

CHANGES IN MEAN READING AND MATH SCORES OVER 15 MONTHS FOR RETURNING COHORT I TREATMENTS AND SAN DIEGO CONTROLS

	Returning Treatments	San Diego		
		Treatments	Controls	Difference
Reading Scores				
1985 Pre-test	715.78	729.93	729.37	
First Summer Change	-5.46**	-12.27*	-13.02***	0.75
School Year Change	18.18***	20.85***	26.40***	-5.55
Second Summer Change	-17.05***	-8.51	-24.80***	16.30*
15-Month Change	-4.33	0.08	-11.43**	11.51
Math Scores				
1985 Pre-test	692.86	733.79	730.35	
First Summer Change	1.34	-12.94	-9.30	-3.64
School Year Change	20.93***	22.30**	22.43***	-0.13
Second Summer Change	-4.07	6.44	-17.41***	23.85**
15-Month Change	18.20***	15.80*	-4.28	20.08*

***Indicates that the impact is statistically different from zero at the 0.01 level.
**Indicates that the impact is statistically different from zero at the 0.05 level.
*Indicates that the impact is statistically different from zero at the 0.10 level.

treatments sometimes show significant losses. One might hypothesize therefore, that test data on the control group in the second summer, if it were available, would indicate losses of a greater magnitude than that observed for treatment youth.

As indicated in Table VII.5, control youth in San Diego experienced significantly greater losses during the second summer of STEP than their treatment counterparts. Treatment youth in San Diego lost 8.51 points (0.5 of a grade) in reading while control youth lost 24.81 points (1.3 grades). The differential performance in math is greater; treatment youth gained 6.44 points (0.3 of a grade), while control youth lost 17.41 points (0.6 of a grade).

The estimated net impact of treatment for San Diego youth during the second summer also indicates the benefit of participating in a second summer of STEP. Regression analysis, which controlled for skill levels measured at the start of the second summer (i.e., 1986 pre-test scores), indicates that treatment youth outperformed control youth by 14.09 points in reading and 21.74 points in math during the second summer. The impact of STEP in San Diego during the second summer is equivalent to 0.7 of a grade in reading and 0.8 of a grade in math, both of which are statistically significant. These impacts are larger than those experienced by San Diego youth during their first summer of STEP.

Although we have no way of knowing whether the impact of STEP in the other four sites was as strong during the second summer as they were in San Diego, it is possible that control group losses at the other four sites would have been larger than those experienced by the treatment youth had we obtained measures for the comparison group.[17]

The large second summer learning losses of the treatments in Cohort I are nonetheless disconcerting. A number of hypotheses have been posed to explain these losses:

- First, we might expect larger second summer losses if the youth who returned for the second summer were those who were least academically proficient among the STEP youth.

- We might also expect that the youth who performed the worst, i.e., who had the largest first summer losses, might be more likely to return.

- Returning Cohort I youth may have had poor attendance during the second summer program and thus received less remediation than Cohort II youth.

---

[17] The impact of STEP in San Diego at the end of summer 1985 was similar to the impacts in the other sites.

- Remediation may have been less effective for returning Cohort I youth than for Cohort II as a result of differences in the quality of instruction and/or differential classroom decorum.

- Finally, attitudes toward the testing situation, particularly at post-testing, may have inflated the magnitude of the observed losses.

With respect to the first hypothesis, the attrition analysis indicated that returnees were not less academically proficient than the original group of treatment youth. As discussed above and in Chapter V, the baseline MAT scores for the two groups were nearly identical. Similarly, the data do not support the hypothesis that youth who lost the most (or learned the least) in the first summer returned. In fact, the first summer loss for the returnees in reading (reported in Table VII.5) is slightly smaller than for the original Cohort I treatments.

The third hypothesis, that Cohort I youth attended fewer hours of remediation than Cohort II youth, is also not supported by the data. As reported in Chapter IV (Table IV.5), classroom hours were nearly identical for the two cohorts. Cohort I youth averaged 99 hours of remediation and life skills instruction, while Cohort II attended an average of 100 hours.

There is no indication, based on classroom observations, that the quality of instruction delivered to Cohort I youth differed from that received by Cohort II. In addition, characteristics of the teachers that were measured (such as experience with individualized instruction or CAI, areas of specialization and reaction to the use of modules) did not differ for the two cohorts.

Classroom observations do suggest, however, that behavioral problems among Cohort I youth may have resulted in less effective treatment compared to Cohort II youth. For example, in Seattle, where remediation was offered in both morning and afternoon sessions, teachers frequently commented that youth in the afternoon sessions were less attentive and more restless. Although Cohort II youth were split between morning and afternoon remediation, all Cohort I returnees attended classes in the afternoon.[18] Classroom observers in other sites also reported more behavioral problems among the returning Cohort I youth in comparison with the entering Cohort II youth.

---

[18] Although the differences were not statistically significant, Cohort II youth in Seattle who attended afternoon remediation experienced larger losses than those in morning classes.

The greater incidence of behavioral problems among the returnees may be due to any of several factors. Simply being a year older than the Cohort II youth may have been one factor. The changes in remediation from the first summer to a more structured program in the second summer meant that Cohort I youth returned to a situation that was quite different than what they had experienced the first summer. Discrepancies between their expectations (based on their first summer experience) and the demands of the revised remediation program could also have increased behavior problems. Although observational data confirmed the existence of behavioral problems among the Cohort I returnees, we have no way of determining what effect these problems may have had on the learning outcomes or the extent to which they would not have occurred were the project to be in equilibrium.

Finally, in two sites (Boston and Fresno), Cohort I youth were tested separately from Cohort II youth. In both of these sites, testing observers noted that Cohort I youth had a less serious attitude toward the test and exhibited more inattentive behavior than Cohort II youth at both the pre-test and post-test. Thus in spite of the precautions taken to reduce measurement error in test scores, it is possible that the large losses experienced by returning Cohort I youth in the second summer are partially due to measurement error.

15-Month Outcomes

Table VII.5 contains changes in mean reading and math scores for returning Cohort I youth from the beginning of STEP (1985 pre-test) to the completion of STEP (1986 post-test). The 15-month program period includes two summers when we would expect youth to experience learning losses, and one school year during which we would anticipate learning gains.

Over the 15 months of STEP, these Cohort I treatment youth experienced a small, non-significant loss in reading, but a statistically significant gain in math. In reading, the gain of 18.18 during the school year is completely offset by the two summers of losses leading to a non-significant loss of 4.33 (equivalent to 0.3 of a grade). In math, however, where the summer losses were smaller, Cohort I treatment youth gained 18.20 points (approximately half a grade) during their STEP enrollment. Over the 15-month program, the estimated marginal impact of each treatment hour is 0.03 points in reading and 0.10 points in math (neither is statistically significant).

The results for San Diego returnees again provide some insight into what the impact of participation over the complete 15-month program might be. As was true for the treatment youth overall, the school-year gain in reading experienced by the San Diego treatments was offset by losses during the two summers. However, while the treatment youth in San Diego maintained their baseline reading level, their control counterparts lost 11.43 points (or 0.6 of a grade), dropping from a 6.5 level at the start of the program to 5.9 at the end. In math, larger summer losses completely offset school-year gains for the controls. However, STEP was able to stem summer math losses for the treatments (at least in the second summer), thus producing a net gain in math over the 15-month period. Since there was no significant difference between treatments and controls in school-year learning, the treatment-control differential for the 15-month program in both reading and math is largely due to differential summer learning.

The net impact of 15 months of STEP for San Diego youth was estimated using a multiple regression model that controlled for 1985 baseline differences between the treatment and control youth. In reading, the impact of the full program was equivalent to 0.6 of a grade (though the estimate is not statistically significant). The net impact on math was equivalent to 0.8 of grade (which is significant at the 0.10 level). As with the second summer outcomes, we cannot generalize these San Diego impacts to the remaining Cohort I treatment youth. We can only infer that the positive impact of STEP in San Diego suggests that participation in the full 15-month program is beneficial in terms of improving basic skills. Given the better first summer performance of the second cohort, we would expect better outcomes during the school year and second summer for Cohort II as well.

Return to School

Participation in STEP did not apparently affect school enrollment. In the fall of 1986, school districts provided enrollment information for the 1986-87 school year for all Cohort I youth. Approximately 90 percent of all Cohort I youth (90.6% of treatments and 88.8% of controls) returned to school in the fall of 1986. Among the Cohort I treatment youth who participated in a second summer of STEP, 94.7 percent returned to school in 1986. In San Diego, 97.3 percent of treatment youth and 94.3 percent of control youth who returned for the second summer of STEP, returned to school. These rates of school enrollment for treatments and controls are not significantly different.

Description of Life Skills Outcomes

Returning Cohort I youth completed questionnaires at the beginning and end of the second summer similar to those administered in the first summer of STEP. Measures of contraceptive knowledge and knowledge of the availability of contraceptives that were identical to those asked in the first summer provide information on the retention of the knowledge gained during the first summer and whether knowledge continued to increase during the second summer. Youth were also asked to provide behavioral information, similar to that requested during the first summer, on sexual experience, recent sexual activity and use of contraceptives. Finally, measures of health and substance abuse knowledge were included in the second summer questionnaires. (These two topics are the subject of individual lessons only in the second summer of the life skills curriculum.)

Again, we are unable to determine program impacts on any of the life skills outcomes since we do not have comparable second summer data for control youth except in San Diego.

Fertility-Related Knowledge

Mean scores on the measures of contraceptive knowledge at all four measuring points for returning Cohort I treatment youth and San Diego returnees are presented in Table VII.6. Looking first at the mean scores for the returning treatment youth, we find that over the complete 15-month program, they significantly increased their knowledge of contraception. As was true for the original Cohort I treatments, returnees significantly increased their knowledge during the first summer on both measures by approximately 50 percent. Not only did youth increase their knowledge during their first summer in STEP, most of that knowledge was retained during the intervening school year. Although the decreases from endline 1985 to baseline 1986 are statistically significant, the losses are very small (about 20 % of the knowledge gained during the first summer was lost over the school year). During the second summer of STEP the level of knowledge increased once again, by about half as much as during the first summer.

When youth entered STEP they were able to correctly answer about five of the 12 contraceptive knowledge questions. By the end of the 15 months, they correctly answered an average of 3.75 more questions. Similarly, youth increased their knowledge of contraceptive availability from about three of the eight questions at baseline to around five and a half of these questions by the end of the program.

The San Diego data provide an indication of the increase in treatments' knowledge relative to the control group's. The San Diego results suggest that the observed knowledge increases are more likely due to participation in STEP than simply to maturation. Although control youth gained some knowledge during the intervening school year (and thus for the overall program period), the knowledge gained by San Diego treatment youth during both summers is significantly greater than for control youth. Over the 15 months, the level of knowledge for San Diego treatment youth increased substantively and significantly more than for control youth.

Table VII.6

CHANGES IN MEAN KNOWLEDGE OF CONTRACEPTIVE METHODS AND AVAILABILITY OVER 15 MONTHS FOR RETURNING COHORT I TREATMENTS AND SAN DIEGO CONTROLS

	Returning	San Diego		
	Treatments	Treatments	Controls	Difference
Mean Knowledge of Contraceptive Methods[a]				
1985 Pre-test	4.70	4.49	5.09	
First Summer Change	+2.58***	+3.10***	-0.01	3.11***
School Year Change	-0.57***	-1.45***	+0.29	-1.74***
Second Summer Change	+1.17***	+2.07***	-0.04	+2.11***
15-Month Change	+3.17***	+3.72***	+0.24	+3.48***
Mean Knowledge of Contraceptive Availability[b]				
1985 Pre-test	2.88	2.77	3.03	
First Summer Change	+2.05***	+2.68***	+0.29*	+2.39***
School Year Change	-0.37***	-1.00***	+0.64***	-1.64***
Second Summer Change	+0.96***	+1.63***	+0.01	+1.62***
15-Month Change	+2.64***	+3.31***	+0.94***	+2.36***

[a] Mean number correct responses on a 12-item scale.

[b] Mean number correct responses on an 8-item scale.

***Indicates that the impact is statistically different from zero at the 0.01 level.
**Indicates that the impact is statistically different from zero at the 0.05 level.
*Indicates that the impact is statistically different from zero at the 0.10 level.

Regression analysis on the San Diego sample provides estimates of the net impacts. During the second summer, STEP increased the youths' scores on knowledge of contraceptive methods by 2.42 questions and their scores on knowledge of the availability of contraception by 1.74 questions. Both of these impacts are statistically significant at the 0.01 level. The net impact for the total 15-month program is 3.04 on knowledge of contraceptive methods and 2.31 on knowledge of contraceptive availability (both estimates are significant at the 0.01 level).

The above discussion suggests that participation in STEP significantly increases these youth's knowledge of birth control beyond what they would have gained as part of the normal maturation process and/or as the result of any sex education received during the intervening year. Additionally, treatment youth retained most of what they learned the first summer over the subsequent year. As one would expect, the San Diego control data indicate that some increase in contraceptive knowledge can be expected over a 15-month period. What treatment youth in San Diego were able to learn in comparison to the controls, however, suggests that STEP is very effective in this area.

Sexual Behavior

Although three indicators are available for analysis of sexual activity during the second summer, we do not have measures of all three at each of the four data collection points.[19] Whether or not a youth has ever had sex was measured at baseline 1985 and at both points in 1986. As discussed above (and as we would expect as these youth mature), the percentage of youth who were sexually experienced increased between the beginning of STEP and the beginning of the second summer. The percentage continued to increase during the second summer of STEP. Among returning treatment youth for whom we have measures of prior sexual experience at both baseline 1985 and endline 1986, the percentage with any sexual experience increased by 25 percent (from 49.1% to 61.4%). The question that remains unanswered at this time is whether this increase is smaller than would have occurred in the absence of STEP.

As with previously discussed program measures for the returning Cohort I youth, we turn to the San Diego data to provide an indication of the relationship between treatment and control youth behavior during this time period. Among San Diego youth, the percentage of treatments with sexual experience increased over the 15 months by 8.2 percentage points, while the percentage among control youth increased by 15.8 percentage points.

---

[19] The data collected in the follow-up interviews will permit a better look at sexual behavior and for a longer period of time than the data currently available.

Although the percentage of control youth with sexual experience increased 7.6 percentage points more than the percentage of treatment youth, we cannot reject the hypothesis that this was a result of chance.

The second indicator of sexual behavior is recent sexual activity, i.e., whether the youth had sex in the two months preceding each measurement. Among the returning treatment youth, approximately 25 percent reported recent sexual activity when they entered the program. This percentage increased significantly over the first summer and the following school year. At the beginning of the second summer, 39 percent reported recent sexual activity. Over the second summer, however, the percentage of youth reporting recent sexual activity stabilized.

The percentage of returning Cohort I youth in San Diego reporting recent sexual activity also increased during the first summer and the intervening school year by about the same amount for treatments and controls. During the second summer, the percentage dropped for both groups with a slightly larger decrease among control youth. None of the differences between treatment and control youth in San Diego are statistically significant. Thus STEP participation appeared to have no effect there.

The final measure of fertility-related behavior is the use of contraception during the last intercourse.[20] Among returning treatment youth the percentage who reported using some form of birth control during their last intercourse increased from the beginning to the end of the second summer by 15 percent (from 35.1% to 40.4%). Unfortunately, with respect to this measure, the San Diego data cannot provide even the rough comparison between treatments and controls that was possible for other outcome measures. Only about 40 treatment and 40 control youth answered the question. These are too few numbers to analyze.

The information available for analysis at this time provides very little insight into the possible impact of STEP on sexual behavior. Sexual activity varies greatly across racial/ethnic groups and between adolescent males and females. Thus the limited information provided by San Diego control youth is rendered even less representative because of the specific racial/ethnic composition of the Cohort I sample in San Diego. Although sexual experience increased over the 15 months of STEP, and contraceptive use reportedly increased during the second summer, the lack of comparable data for the control group precludes drawing any conclusions about impact.

---

[20] Due to changes from the first to the second summer in the presentation of this measure in the questionnaires, the information regarding use of contraceptives is not comparable for the first and second summer. Thus we only examine changes during the second summer.

Extensive fertility-related information, including sexual activity, contraceptive use, pregnancies and pregnancy outcomes will be collected in the future as part of the follow-up interviews with both treatment and control youth. These data will enable us to determine whether any of these outcomes were affected by participation in STEP.

Health and Substance Abuse Outcomes

Two topics that were included as individual lessons in the second summer's life skills and opportunities curriculum that were not part of the first summer's curriculum were general health issues and substance abuse. While Cohort I treatments' knowledge about general health issues did not change over the summer, their knowledge about drug and alcohol abuse increased.

To measure whether knowledge about health increased during the summer, youth were asked to respond to four true-false statements on the pre-test and post-test questionnaires. As shown in Table VII.7, at the beginning of the summer, treatment youth correctly answered an average of 2.62 of these questions. The level of knowledge as measured by these statements did not change over the summer. In addition, among the San Diego returning youth, this treatment change of 0.12 and the control change of -0.11 on this scale were not significant, nor was the treatment-control difference. Thus in San Diego, participation in STEP did not affect knowledge of general health issues.

Knowledge of substance abuse was measured by eight true-false questions. Returning treatment youth correctly answered 4.27 of the eight questions at baseline. Their knowledge increased significantly over the summer by 0.53 questions. Treatment youth in San Diego increased their knowledge over the summer by 0.89 questions (which is significant at the 0.01 level). The increase in knowledge among San Diego treatment youth was significantly more than the 0.13 additional questions answered correctly by control youth. Thus, for youth in San Diego, the LSO session on substance abuse significantly affected knowledge of this topic.

Summary

In summary, Cohort I treatment youth who returned for a second summer of STEP experienced a significant loss in reading and a small, non-significant loss in math during the second summer. Over the 15-month program period, returning treatment youth experienced a small, non-significant loss in reading and a substantial gain in math.

Table VII.7

BASELINE TO ENDLINE CHANGES IN MEAN HEALTH KNOWLEDGE AND MEAN DRUG KNOWLEDGE FOR COHORT I

	Returning Treatments	San Diego Treatments	San Diego Controls	Difference
Mean Health Knowledge[a]				
Pre-test	2.62	2.59	2.91	
Post-test	2.68	2.71	2.81	
Change	0.06	0.12	-0.10	0.22
Mean Drug Knowledge[b]				
Pre-test	4.27	4.14	4.15	
Post-test	4.80	5.03	4.28	
Change	0.53***	0.89***	0.13	0.76***

[a] Mean number of correct responses on a 4-item battery.
[b] Mean number of correct responses on a 8-item battery.

***Indicates that the impact is statistically different from zero at the 0.01 level.
**Indicates that the impact is statistically different from zero at the 0.05 level.
*Indicates that the impact is statistically different from zero at the 0.10 level.

The San Diego sample, which included control youth who returned for a second summer job as well as returning treatment youth, suggests that STEP successfully stemmed summer learning losses in that site. In spite of nearly equivalent first summer losses and school-year gains, control youth in San Diego lost significantly more than treatment youth in both reading and math during the second summer.

Cohort I treatment youth who returned for the second summer continued to increase their contraceptive knowledge. There is no evidence at this time, however, of an impact on sexual behavior. The information collected in the first wave of follow-up interviews will be better suited to uncovering any impact on sexual behavior for Cohort I youth.

VIII. COST ANALYSIS

This chapter examines the costs involved in operating the 15-month Summer Training and Education Program that are above and beyond the costs of a program providing only work experience to a similar group of 14- and 15-year-old youth. First, to lay out the framework, we discuss the comparison group used for this analysis, explain how 15-month cost estimates are derived from 12 months of cost data, describe the perspective from which costs are examined, present several complications arising in the cost analysis and how they were handled, and identify the data sources.

Examination of demonstration costs indicates that if STEP were replicated, the incremental social costs of sending a student through a "steady-state" 15-month summer training and work experince program like STEP, instead of providing them work experience only, would be approximately $1,600 per enrollee (assuming such a program enrolled 150 new students per year). In the following sections we develop this cost estimate component by component, examining the costs incurred in each site and the cost variation among the sites. Each component cost section concludes with our best estimate of the social costs of replicating STEP. It should be noted that this is the first year that sites reported costs on STEP and these data may be subject to considerable measurement error. Thus, it is our belief that next summer's cost estimates will be measured more precisely and be more realistic.

METHODOLOGY

While an analysis of program costs is in some respects straightforward, several conceptual issues arise in providing accurate estimates, as follows.

Comparison Group

The costs presented in this chapter are the costs over and above providing work experience to 14- and 15-year-old youth. Thus, the comparison program used for the cost analysis is consistent with the control group used for the impact analyses, with the exception that the comparison program (SYETP) does not include an educational eligibility criterion. Thus, the efforts made to identify whether students were having difficulty in school were counted as part of net STEP program costs.[1]

---

[1] Legislation passed after the cost analysis plan was developed requires SYETP to use some method to determine a student's need for remediation. The costs of an alternative

Synthetic (Hypothetical) Cohort Technique

The major issue in the STEP cost analysis is determining the cost of providing a full 15 months of STEP service to youth. To address this issue, sites submitted to P/PV the costs they incurred during the months between the end of summer 1985 and the end of summer 1986.[2] During these 12 months, the program provided school-year support for Cohort I treatment youth, and offered summer training and work experience to treatments from Cohort II (who were enrolled for the first time) and Cohort I treatments (who were re-enrolled). These data do not reflect the full cost of sending youth through the 15-month program. Nevertheless, by using a synthetic (hypothetical) cohort approach, one can derive an estimate of the cost of the full program. The component costs of STEP and the sources from which they are derived are found in Table VIII.1.

Knowing cost by component, the number of Cohort I youth, the number of Cohort II youth, and the percentage of youth who returned for a second summer, we constructed the cost of 15 months of STEP for a hypothetical cohort that received two summers of training (as it was delivered in 1986) and school-year support (as it was provided during school year 1985-86).

Cost Perspective

The costs of a program such as STEP, as well as the benefits, can be analyzed from a variety of perspectives--the government's, the participant's, or society's. The most common approach in benefit cost evaluations--and the one many would argue is the most appropriate--is to analyze program costs from the perspective of society as a whole. This approach focuses on the overall value of goods and services invested by society in the program. From the societal perspective, changes in the ownership of resources by groups within the society (transfer payments)--such as payments to STEP participants--are not counted as costs. They represent one group's costs (the government's) that correspond to and exactly offset benefits to another group (the participants). Only changes in the total value of resources are counted as costs from society's perspective, costs such as the operational

---

program may well include educational eligibility costs, and thus, reduce the net cost of STEP. However, because educational eligibility costs were small ($3 per enrollee), a comparison group incorporating the required eligibility criterion would change our cost estimate little.

[2] These are presented in Appendix D.

Table VIII.1

THE SYNTHETIC COHORT APPROACH: SOURCES OF COMPONENT COSTS

Components	Source
Original Eligibility Determination	Cohort II in 1986
First Summer Costs (Remediation, LSO and Work Experience)	Cohort II in 1986
School-Year Support	Cohort I in school year 1985-86
Re-Enrollment	Cohort I in 1986
Second Summer Costs (Remediation, LSO and Work Experience)	Cohort I in 1986
Administrative Costs	Cohorts I and II in 1986

expenses involved in making transfers. Because the societal perspective represents the appropriate approach for estimating the economic efficiency of the program, it is used in this analysis.

Cost Evaluation Issues

Before tallying costs, a number of issues that affected the final cost estimates had to be resolved. These included defining the steady-state costs, disentangling operational versus research costs, valuing contributed goods and services, and treating joint costs. We discuss each of these below.

Costs of Steady-state Operations

Establishing the program incurred considerable development costs, particularly those for creating the remediation and LSO curricula. However, the amortized value of these expenditures has not been included in the estimates discussed here. The cost analysis estimates the costs of the program during its steady-state operation because these represent the long-run costs. Yet, in demonstrations, relatively large proportions of the costs incurred are either build-up or phase-down costs. Thus, cost estimates based on demonstration experiences must be careful to exclude these costs.

There are two methods to determine the cost of a program during its steady state. Either start-up and phase-down costs are quantified and omitted; or costs are observed during a period following initial start-up but before beginning the phase-down process. The latter approach is used here; the cost analysis covers costs incurred during 1986. In all sites except Boston (which was also a pilot site), this period covers the second summer of program operations. We are reasonably confident that the costs of start-up activities, for the most part, are not included in the analysis.

Curriculum development in 1986, however, entailed extensive start-up costs. At the end of the summer of 1985, sites requested that P/PV provide them with a standard curriculum that they could tailor and modify to their needs. Though some year-to-year curriculum changes are likely to occur in a steady-state program, the intensity of the curriculum development activity that occurred in 1986 is greater than what would occur in steady-state. A more reasonable approximation of steady-state costs would be the curriculum revision costs incurred in 1986-87. Although P/PV's 1986-87 costs were available, curriculum development activities for 1987 had not yet taken place in the sites. Thus, the curriculum development costs presented here are the 1986 costs for the sites and the 1987 costs for P/PV.

While the cost estimate for curriculum development may overestimate the steady-state cost, the cost of school-year support for 1985-86 is likely to underestimate it. The school-year support program, as it was run in 1986, was not as strong as it would probably be in steady state. In 1987, efforts were made to strengthen the program and, thus, the cost of the program is likely to increase from its 1985-86 level.

Operational Versus Research/Demonstration Costs

Another common problem in evaluating costs of a demonstration like STEP is the exclusion from internal operational costs, those costs associated with the demonstration's research requirements. These are costs incurred because the intervention is the focus of special research inquiry or because more intensive efforts are needed to implement an intervention that enables us to extrapolate findings to other populations. These costs are called demonstration costs; the cost of providing jobs and testing the control group are examples.

The assignment to operations of P/PV's local site coordinators' salaries is not as straightforward. For the most part, these payments are demonstration costs because the primary functions of the coordinators were to monitor and document the implementation of the program, clearly research activities. But the coordinators also provided technical assistance to ensure the smooth operation of the program. Even in such cases, however, it could be argued that these activities represented demonstration costs. The suggestions of the coordinators were made, primarily, to ensure the site's conformity with the model. However, in particular cases, we determined that the involvement of site coordinators went beyond their research roles. As a general rule, we took a conservative approach to this assignment process, considering site coordinator costs to be operational only if they were clearly and unmistakably beyond what would have been expected as part of the demonstration.

Actual Versus "Contributed" Goods and Services

An analysis of program costs from society's perspective must consider the value of goods and services that are contributed, in addition to those paid for from program funds. They represent important program resources but are frequently overlooked or misestimated because they may not appear in the formal accounting records. Yet, had the resources not been used for program purposes, society could have used them for other purposes (the opportunity cost).

While the concept of contributed goods and services is straightforward, the estimation of their value is more difficult. They are sometimes elusive, especially when the costs are identified retrospectively, as is the case in this analysis; respondents have a more difficult time recalling resources that are not formally a part of the program budget. Another problem is determining the opportunity cost of the resource, since its alternative use may be in question. For example, donated classroom space and computers might have sat idle for the summer had they not been used by STEP. Even so, we have elected to include estimates of these costs to make them comparable across sites--one site had to rent some computers. It should be noted that their exclusion would not have had a major impact on the cost estimates since the primary program expenditures are for teachers' salaries.

Treatment of Joint Costs

Another problem faced in most cost analyses is the appropriate treatment of joint costs, which pertain to more than one program or program component. Joint program costs are the more serious because, unless they are allocated appropriately, they might result in an under- or over-estimation of STEP costs. An example of joint program costs is the funding of services provided both to STEP participants and control group members who were SYETP enrollees. Joint costs (such as job development) were incurred in a few instances and were allocated appropriately.

Separation of costs that pertain to more than one component of a program is not as crucial for determining overall program cost. Determining the cost of each component is, however, difficult. The separation of program administration costs from the costs of delivering remediation and life skills was particularly problematic.

Data Sources and Issues

The primary sources of data were the accounting and other program records kept by the lead agency at each STEP site, as well as by other agencies (such as school districts) that were involved in operations. Another important source of information was the staff of each of these organizations. In each of the sites, the P/PV coordinator was responsible for identifying the appropriate source of records and respondents in each organization, obtaining the needed information for the relevant time period, and conducting a preliminary analysis of the information to determine its accuracy and need for adjustment or revision.

It should be noted that neither site coordinators nor local agency staff were required to maintain detailed records on program costs during the period of program operation analyzed here.[3] Thus, the data collection effort was retrospective and constrained by the extent of information that had been retained. We are making efforts to remedy this in the coming year.

---

[3] Sites submitted budgets for the funds P/PV provided to them. However, because these were not the only resources used in STEP, these budgets are not sufficient for the cost analysis.

THE COSTS OF STEP

This section presents the costs (both labor and non-labor combined) for each major component and explores variations by site. Costs are presented as per-enrollee costs,[4] as opposed to total costs. This adjusts for differences in program sizes across the sites. Table VIII.2 displays first summer, second summer and school-year costs separately. For the components that are offered in both summers, such as remediation, the costs are assumed to be the same for the two summers. Thus, the site variations are discussed in terms of the first summer costs. The discussion of each component ends with our best estimate of the cost of replicating STEP. The section ends with a discussion of cost estimates for the 15-month program.

Remediation

The costs of remediation in 1986 are separated into two categories: summer instruction and curriculum development. We discuss each of these categories below.

Summer Remediation Instruction

The major cost of operating STEP, over and above running a work experience program, was the cost of instruction. Considering the costs in all the sites and adding those incurred by P/PV to train teachers and provide technical assistance, the total instructional cost per enrollee was approximately $572 in the first year and $572 in the second year for those who returned for a second summer, as shown in Table VIII.2. (No second-year costs were incurred for the 25 percent who did not return.) The site costs during 1986, upon which these estimates are based, varied widely; they ranged from Fresno's cost of $402 per enrollee to San Diego's $818.

---

[4] A third cost estimate that could have been considered was cost per slot. This is the cost per hour of actual participation multiplied by the number of hours that composes full-time participation. In one sense, this estimate is more meaningful than cost per enrollee because it removes individuals who were absent much of the summer. However, slots are not filled when a participant drops out, thus, we would never expect full-time participation. In addition, cost per enrollee is the appropriate cost measure to compare with program impacts, which are also on a per-enrollee basis.

Table VIII.2

ESTIMATED COSTS PER COHORT MEMBER BY SITE FOR 15 MONTHS OF STEP

Program Component	Boston	Fresno	Portland	San Diego	Seattle	Demonstration Average (Including P/PV Costs)	Replication Estimate
First Summer Costs							
Remediation							
Summer Instruction	$513	$402	$473	$818	$599	$572	$572[a]
Curriculum Development	90	69	27	30	47	70	30[b]
Total	603	471	500	848	646	642	602[c]
Life Skills & Opportunities	45	34	49	54	62	62	62
Work Experience	0	3	14	0	7	5	5
Other Activities & Administration							
Eligibility Determination	3	1	0	0	12	3	3
Program Administration and Support Services	10	23	89	52	58	45	45
Total	13	24	89	52	70	48	48
Total First Summer Costs	661	532	652	954	785	757	717
School-Year Support Costs	128	301	196	0	168	198	300

Table VIII.2 - (cont'd)

ESTIMATED COSTS PER COHORT MEMBER BY SITE
FOR 15 MONTHS OF STEP

Program Component	Boston	Fresno	Portland	San Diego	Seattle	Demonstration Average (Including P/PV Costs)	Replication Estimate
Second Summer Costs							
Remediation							
Summer Instruction	$513	$402	$473	$818	$599	$572	$572[a]
Curriculum Development	90	69	27	30	47	70	30[b]
Total	603	471	500	848	646	642	602[c]
Life Skills & Opportunities	45	34	49	54	62	62	62
Work Experience	0	3	14	0	7	5	5
Other Activities & Administration							
Eligibility Determination	75	17	115	na	58	66	66
Program Administration and Support Services	10	23	89	52	58	45	45
Total	85	40	204	118[d]	116	111	111
Total Second Summer Costs	733	548	767	1,020	831	820	780
Expected 15-Month Costs Per Original Enrollee[e]	$1,243	$1,283	$1,510	$1,674	$1,602	$1,587	$1,619

NOTE: The component totals may not sum to the overall totals because all entries were rounded to the nearest dollar.

[a] These estimates include $17,200 of P/PV incurred costs.

[b] These estimates include $23,250 of P/PV incurred costs.

[c] These estimates include $40,460 of P/PV incurred costs.

[d] We have assumed that San Diego would have had average re-enrollment costs ($66).

[e] Second summer costs are incurred for those youth who return for a second summer. Thus, expected 15-month costs, averaged over the first summer enrollees, is not the sum of the first and second summer plus school-year support costs but the sum of first summer, school-year support and re-enrollment costs plus other second summer costs in proportion to the percent of youth who return. The proportion of youth who returned varied from approximately 58 percent in Boston to 84 percent in Portland. The demonstration average was approximately 75 percent. Do you ever have problems understanding a teacher when he or

In all sites except San Diego, the majority of the costs were labor costs--instructor salaries and benefits. During summer 1986, labor costs (not broken out in Table VIII.2) ranged from $231 per enrollee in Fresno to $441 in Seattle, depending on the number of instructors (and thus the class size) and the teacher salaries and benefit rates. Teacher salaries were governed by local collective bargaining agreements and were out of the control of the program. For example, the average wage in San Diego was higher than at other sites, in part, because STEP teachers were all regular school district personnel paid their regular salary with full fringe benefits. The fact that Seattle hired the fewest remediation teachers did not reduce their costs because these teachers worked both morning and afternoon and had generally higher salaries than teachers in other sites.

The other component of summer instruction costs is non-labor cost, primarily for computers and space, but including costs for resource materials, printing and transportation. Although computers and space were frequently donated by the school districts, the use of these resources represents a cost from the social perspective. Non-labor costs were fairly similar across four of the five sites in 1986, ranging from $166 per enrollee in Boston to $220 in Portland.

San Diego's per enrollee non-labor cost was $465. The non-labor cost in San Diego was dominated by the imputed costs of donated space and computers. Because instruction took place at five locations, many classrooms were used.[5] The extensive use of space represents a considerable cost to society, when valued. If San Diego had had to pay for this space, it probably would not have used as much as it did. Similarly in San Diego, participants used a large number of personal computers.[6] When valued, this program feature is fairly expensive.

Curriculum Development

As explained earlier, the curriculum development costs used here reflect what the sites spent in 1986 and what P/PV spent in 1987 when start-up costs were reduced. Curriculum development activities at the sites during 1986 probably cause the 1986 costs to overstate true steady-state costs. In particular, each of the sites spent resources to develop a local module and customize the P/PV modules. Steady-state program operations would probably include some curriculum revision costs, but it is difficult to estimate their magnitude.

---

[5] The estimate of the rental value of the classrooms provided by the school district was $75 per classroom per day. This rental price seemed too high relative to the estimates provided by the other sites, thus the average of the rental price in the other sites was used in its place to construct the cost estimate. However, even with this adjustment, space costs are high because of the large number of classrooms used.

[6] Computers and printers, numbering 300, were available in the five schools.

Table VIII.2 shows that average curriculum development costs in the demonstration during a summer are $70 per enrollee. This figure includes $23,250 spent by P/PV on curriculum modification and printing in 1987, as well as the resources spent in 1986 by the sites. The costs varied widely. Boston's cost was relatively high ($90) because that site spent more time and used more highly paid individuals than other sites in developing its own module and customizing P/PV's.[7]

The average total cost of providing one summer of remediation (including curriculum development and summer instruction) for the demonstration is $642 per enrollee. Our best estimate of the replication cost, however, is $602 per enrollee. The replication estimate is lower than the demonstration average because we assume that curriculum development costs would be lower in a steady-state program.[8]

Life Skills and Opportunities (LSO)

The cost of the LSO component of STEP was fairly small, averaging $62 per enrollee per summer. This primarily consists of the LSO teachers' salaries. The $62 total may underestimate actual LSO costs because sites had difficulty dividing costs between the remediation and the LSO components. The yearly averages account for the $18,501 spent by P/PV (in 1987) that was considered to be operational, such as the costs of LSO curriculum revision, training, technical assistance and materials.[9] The sites' costs ranged from Fresno's $34 per enrollee to Seattle's $62.

---

[7] Substantial amounts of time of a senior SYETP official and the P/PV coordinator were used in these activities.

[8] We use the demonstration average cost of summer instruction, without excluding San Diego's very high costs, because if STEP were to be replicated nationally, large numbers of computers and large amounts of classroom space would similarly be donated in some sites.

[9] Just under $3,000 was spent for LSO curriculum revision (specifically Volume I of the LSO curriculum), approximately $5,000 for training, $6,600 for training-related travel, $2,300 for technical assistance and $1,800 for technical assistance training. P/PV's cost of developing Volume II are considered to be start-up costs and were not included in the estimate.

Our best estimate of the cost per participant for one summer of LSO is the demonstration average of $62.

Work Experience

The approach taken in assessing the costs and benefits of STEP in this report is to consider STEP relative to a program that provided 14- and 15-year-old youth with work experience alone. Thus, the work experience cost estimate includes only the net costs of providing STEP treatment youth with part-time, as opposed to full-time, jobs. Three sites--Fresno, Portland and Seattle--indicated that they incurred net work experience costs during 1986. In Fresno the per enrollee cost was $3, in Portland it was $14 and in Seattle it was $7. Staff members attributed these extra costs to developing new part-time jobs. In steady state, when employers who desire part-time help have been identified, these costs would probably become negligible in all sites, as it was in two sites in 1986.

Our best estimate of the incremental cost of a STEP job slot (as opposed to a regular SYETP slot) is the demonstration average of $5 per enrollee per summer.

Other Services and Administration

This category encompasses enrollment activities, (including eligibility determination) program administration and several functions ancillary to the primary intervention.

In the first year, one these costs--enrollment activities--was associated with incorporating the "academic level" criterion into overall eligibility requirements. All but one site reported costs under $5, which may be a substantial under-estimate. Seattle reported costs of $12 per enrollee, incurred as a result of relatively poor cooperation between STEP and school district staff.

Included in this component are costs associated with hiring and managing the summer school teachers, payroll, enforcement of STEP rules (in particular, "no school, no work"), and support services, such as general counseling and transporting students to the work place or for special field trips. The average across the sites was $45 per enrollee, but this varied considerably by site. Boston reported only $10 of administrative cost, clearly an underestimate; administrative costs were not well separated from other costs. Portland, on the other hand, reported administrative costs of $89 per enrollee, caused primarily by payroll system problems. Gathering and recording the hours spent in work, remediation and LSO proved to be more time consuming in Portland than in the other sites, probably because a new computerized payroll process was introduced.

When the costs of all these activities were summed, total administration and support costs over one summer varied from $13 per enrollee in Boston to $89 in Portland. The demonstration average was $48 per enrollee. Our best estimate of general program administration costs is the average of $48 per participant, assuming that the program enrolled approximately 150 new participants each year.

Re-enrollment replaces eligibility determination in the second summer. The costs of re-enrolling the Cohort I treatment group for its second summer varied greatly by site, ranging from $17 in Fresno to $115 in Portland.[10] The variation in the costs is almost entirely related to the variation in efforts required to contact students and convince them to return. There were small variations in the average salaries of staff used to contact the youth.

School-Year Support

All sites were provided with grants from P/PV to maintain contact with STEP participants and to operate a school-year support program. However, efforts in each site during school year 1985-86 varied considerably. San Diego did not conduct school-year support activities, while the Fresno and Portland programs were fairly active. The per enrollee costs in the four sites ranged from $301 in Fresno to $128 in Boston. Fresno's costs were the highest not only because it was the most active, but also because a school district administrator and an advisory board became involved.

School-year support was much weaker in the 1985-86 school year than it is in the 1986-87 school year or than it would be if STEP were replicated. Our best cost estimate of future school-year support programs is, therefore, higher than the demonstration average. The total amount of money distributed by P/PV for school-year support in 1986-87 was 50 percent larger than in the year before. Thus, to approximate the cost of school-year support, we inflate the demonstration average of $200 by 50 percent to get a cost estimate of $300 per cohort member.

TOTAL 15-MONTH PROGRAM COSTS

The estimated net cost of a youth enrolled in the 15-month STEP program in steady-state, over the cost of providing work experience alone for two summers, was $1,587. This includes the costs of the first summer, school-year support and a pro-rated amount of second summer costs. The full second summer costs are not included because they are incurred only for those youth who return.[11] Three of the sites--Boston, Fresno and Portland--had total net costs per enrollee of between $1,200 and $1,500. San Diego and Seattle had per enrollee net costs of between $1,600 and $1,700 (despite the lack of a school-year support program in San Diego). Our best net cost estimate of operating a STEP intervention that is in steady state and includes an active school-year support program is approximately $1,600 per enrollee (for sites enrolling approximately 150 new individuals per year).

---

[10] Re-enrollment costs for San Diego are not comparable to those in the other sites. San Diego was the only site that re-enrolled controls. They reported only the extra cost of re-enrolling treatments about that of re-enrolling controls.

The total net cost estimate above does not include the money spent to pay the youth for their classroom and work hours. Such payments are transfer payments--benefits to participants, but exactly offsetting costs to the government--and do not constitute a net cost to society. From an operational perspective, however, the total cost of including these transfer payments is relevant. We estimate that the payment per enrollee is approximately $800 in the first year. Those who return for a second summer (75% of the cohort), again are paid an average of $800; those who do not return are paid nothing. Thus, on average, $600 per original enrollee is paid in the second year.[12] The addition of these transfer payments to the net cost of STEP brings the total cost to approximately $3,000 per enrollee.

It should also be noted that the cost estimates reported in this chapter are from the social perspective and may not reflect the actual expenditures a STEP site would incur. The additional cost to a site for participant wages has just been discussed. Included in the social costs of STEP are contributed goods and services such as classroom space and computers. While these are goods that are integral to the program, from a site's perspective they do not represent real costs if they are contributed. By arranging for such materials to be contributed rather than purchased or rented, a site's actual costs would be lower than those reported here.

Another cost estimate of policy interest is the cost of operating a one-summer STEP program. Our cost analysis indicates that the net costs of such a program (those that exceed the cost of operating SYETP) would be $717 per enrollee, not counting the transfer payments of $800. Included in this estimate are the net costs of eligibility determination ($3); one summer of remediation ($572), LSO ($62) and work experience ($5); and that portion of curriculum development and administration attributable to the first summer ($75). It excludes school-year support (which would add $300 per enrollee), all second summer expenses and re-enrollment costs.

---

[11] The prorated cost per enrollee in each site for the second summer is derived by multiplying the first summer cost by the proportion of the cohort who returned for a second summer in that site.

[12] On average, both Cohort II enrollees and Cohort I returnees were paid for 205 hours of participation. We value these hours at the minimum wage of $3.35 and add 15 percent for administrative costs.

While a one-summer program is substantially less expensive than a two-summer program, it would also be substantially weaker as an intervention. It is difficult for us to assess the effectiveness of a one-summer program relative to a two-summer program, because we do not have follow-up on a one-summer program. The results from the analysis conducted on the 1985-86 school outcomes for Cohort I, presented in Chapter VII, indicate that one summer of STEP with modest effects on basic skills--followed by a school year of meager school-year support--produced little measurable intermediate effect on school performance. However, in next year's report we will be able to determine the impact of one summer of STEP coupled with stronger school-year support (at an average cost of $1,017 per enrollee--$717 plus $300) on the school performance of Cohort II youth during the intervening 1986-87 school year.

IX. SUMMARY AND CONCLUSIONS

The Summer Training and Education Program (STEP) is a five-site demonstration program offering two summers of remediation, life skills instruction and work experience to a randomly assigned group of economically and academically disadvantaged 14- and 15-year-olds. This report documents the operation and the impacts of the program as of the end of the summer of 1986. It outlines the implementation of STEP in its second full summer of operation, when Cohort II youth were served for the first time and returning members of Cohort I received a second (and final) summer of STEP services. The impacts reported are those observable by the end of the summer. For Cohort I youth, we also examine outcomes at the end of school year 1985-86, just before their second summer of participation. For Cohort II, we examine outcomes after the first summer of participation. The cost of STEP during the 15-month program is also estimated. This chapter summarizes the report and discusses the implications of the findings for policy and program development.

STEP IMPLEMENTATION DURING THE SECOND SUMMER

STEP operated more efficiently in the second summer of the demonstration than it did in the first. The improved implementation was due primarily to five factors: greater collaboration between participating school districts and SYETP agencies, earlier planning for the summer program, fundamental changes and improvements in the remediation curriculum, improved management and structure of the remediation component overall, and solution of the start-up problems experienced by sites during the first year.

Early Planning and Improved Collaboration

School districts provided lists of youth potentially eligible for the 1986 cohort recruitment effort earlier than was the case in Year One. This made it possible for sites to complete JTPA eligibility checks earlier for 14- and 15-year-old applicants, and facilitated the cross match with academic eligibility lists. School staff were more directly involved in identifying and encouraging youth to apply for the program. Although these factors expedited enrolling of the new cohort, re-enrolling Cohort I youth for the second summer was a new and different undertaking. Many Cohort I youth needed to be contacted several times before they would agree to re-enroll.

Remediation

At the end of the first summer of operation, the sites declared that STEP remediation could be improved if P/PV provided a curriculum to be tailored and adapted to local needs. Thus, before summer 1986, P/PV developed such a curriculum, called Practical Academics. It offers basic skills instruction by using topic-based modules of interest to 14- and 15-year-olds. From a research point of view, the more standard curriculum produced remediation that was more uniformly delivered across sites.

Other changes that improved the remediation component included: hiring a lead teacher in each site to plan and manage the component; improving teacher selection and training; adding a silent sustained reading element; improving computer-assisted instruction; and increasing technical assistance to sites. Teachers liked the module approach to instruction and, in general, found the philosophy of the STEP curriculum to be compatible with their own.

Life Skills and Opportunities (LSO)

LSO continued to enjoy a high level of success in all STEP sites during Year Two. LSO changes in the second year were: introduction of a new curriculum (Volume II) for returning youth and minor modifications to the first volume. As in Year One, Volume I was very well-received by both instructors and students. Volume II, which covers health, substance abuse and the world of work, was well-received but with less enthusiasm than Volume I. Modifications for Year Three make the second volume more motivating.

Work Experience

This component was also well implemented during the second year. For the most part, returning youth were placed in upgraded jobs; a wider variety of work sites and job types were obtained for youth during 1986 than was the case in the summer of 1985.

School-Year Support (SYS)

The school-year support component, offered for the first time during the 1985-86 academic year, met with the least amount of success in Year One. In the four sites where school-year support was offered, attendance at activities was low, and follow-up with youth was not as organized and effective as P/PV had intended.

Therefore, P/PV provided more programmatic direction and resources during the 1986-87 school year. A half-time coordinator for school-year support was hired in each site. In addition, trained mentors or counselors are required to meet regularly with assigned caseloads of youth, and small and large group activities are being provided.

PARTICIPANTS

A second cohort of youth joined STEP in the 1986 summer. Based on their demographic and economic profiles, Cohort II youth fit the description of youth most likely to drop out of high school. They are mostly minority, poor (JTPA eligible), behind in school and from female-headed households. The educational deficiencies of these youth indicate the need for an intervention such as STEP. Cohort II youth tested approximately two grades below the appropriate levels for their age/grade in reading and approximately one and one-half grade levels below in math. Further, the relative ignorance of this population (42% of whom reported sexual activity) about contraceptives suggests that STEP could potentially affect fertility-related behavior. These youth are similar in most respects to the first group of STEP youth (Cohort I). However, Cohort II is a more racially and ethnically diverse group--participation by Hispanics and Asians increased over Year One of the demonstration.

Approximately 75 percent of Cohort I youth who were asked to return for their second summer of STEP did so. Controls were invited to return in San Diego. For the most part, the characteristics of the returnees mirror those of the original group served. Across all sites, grade level and racial/ethnic background was correlated with the decision to return--re-enrolling youth were likely to be in lower grades and from minority groups. Younger youth were also more likely to return, as were youth who had no work experience or sexual experience prior to joining STEP. This is consistent with the notion that older, more job-ready youth had other opportunities beyond STEP.

Program participation was high among both cohorts of youth. On average, treatments participated in approximately 82 percent of available classroom hours and about 80 percent of work experience hours offered. Participation by controls was slightly higher; they were present about 83 percent of the available time. Consistent with the first program year, the best predictor of level of participation was school-year attendance. Youth with the best attendance during the school year were likely to have the best attendance during the summer.

SUMMER IMPACTS

The research indicates that STEP had substantively large effects on the second cohort's basic skills competencies, knowledge of contraception and sexually-related behavior during the first summer of participation.

Control group learning losses have been large for all years of the demonstration. These losses, experienced by educationally and economically disadvantaged youth in STEP, range from 0.7 to

1.0 of a grade equivalent in reading and between 0.4 and 0.5 of a grade equivalent in math. These losses are large and suggest that programs like STEP that are designed to counteract summer learning losses are worthy of consideration.

STEP's effects on the basic skills of Cohort II participants during their first summer of participation were more than double those of the first year. The net effect was 0.6 of a grade equivalent in reading; where most of the learning loss was stemmed. The net effect was 0.8 of a grade equivalent in math where the loss was not only stemmed but a small gain was produced. These larger results likely resulted from the improved design and implementation of the STEP curriculum.

The effects on Cohort II's basic skills were more widespread than those for Cohort I's first summer of STEP. In math, treatment youth of both sexes, all racial/ethnic groups and all five demonstration sites benefited from participation in STEP. In reading, both sexes, black youth and youth in four of the five demonstration sites benefited significantly.

In summary, the learning effects were larger and more pervasive for this cohort. These effects are more likely to result in the desired longer-term outcomes of school retention and graduation than are the more modest first summer effects found for Cohort I.

STEP's effects on Cohort II's sexual knowledge, attitudes and behaviors were also quite strong. The two scales of knowledge of contraceptives and their availability were very much higher; the mean baseline level increased by over 50 percent. These effects, as well as those on knowledge of the consequences of adolescent pregnancy, were found for both sexes, all racial/ethnic groups and across all demonstration sites.[1]

A priori, we did not expect to see effects on sexual behavior over the six-week period, since the curriculum sessions that most directly relate to sexual behavior (e.g., role playing on how to say no, instruction on physiology and contraception, and the visit to the health clinic) are not delivered until mid- or late summer. Nevertheless, some behavioral effects were demonstrated. STEP had a large effect on contraceptive use by its sexually active participants, increasing the probability of usage by 53 percent overall.

We are not able to discern what effect, if any, STEP had on sexual activity over the summer. The application of two different measures produced conflicting results. Nevertheless, it is encouraging to find that such a brief intervention resulted in some documented effects on contraceptive behavior.

---

[1] This is true with one exception: youth in the whites and others category did not significantly increase their knowledge of the consequences of adolescent pregnancy.

SCHOOL-YEAR IMPACTS FOR COHORT I

We examined the effect that one summer's participation in STEP had on Cohort I's subsequent school behavior and performance, with particular attention on: attendance, standardized test scores, credits earned, grade promotion and dropout behavior. Overall, the only school outcome affected by STEP was grade promotion: treatment group members were 22 percent more likely to pass than were controls. This is a fairly large impact, given that about 24 percent of Cohort I control youth failed. It appears that the program had an especially large impact on promotion rates for Hispanics. Although the impact on promotion is encouraging, we do not correspondingly find that treatment youth earned more credits. Thus, we are cautious about drawing strong conclusions from this one positive result.

The lack of strong effects on school outcomes is not surprising given that the impact of STEP was small for Cohort I at the end of its first summer. At that time, treatments scored only a quarter of a grade equivalent greater in math and reading than controls. Perhaps stronger impacts (such as we see for Cohort II), a second summer of STEP participation, and strong school-year support will produce more substantial effects on school outcomes. We will be able to investigate these possibilities after school year 1986-87 is completed.

SECOND SUMMER RESULTS FOR COHORT I

Since second summer data for control group youth in Cohort I is not available, we can only describe outcomes and not impacts, when considering what happened to Cohort I during the second summer of STEP. While academic and sexual outcomes for all treatments and controls will be obtained through school record and follow-up interview data, basic skills and questionnaire data were obtained from controls in only the first summer of STEP, except in San Diego where controls were promised a job and tested in the second summer. Therefore, net impacts for the second summer cannot be derived, except in San Diego.

During the second summer, Cohort I treatment youth who returned to STEP experienced a significant loss in reading and a small, non-significant loss in math. Over the 15-month program period, these youth experienced a small, non-significant loss in reading and a substantial gain in math.

Analysis of the San Diego sample, which includes both treatment and control youth, suggests that STEP successfully stemmed summer learning losses there. In spite of nearly equivalent first-summer losses and school-year gains, control youth in San Diego lost significantly more than treatment youth in both reading and math during the second summer.

Returning STEP youth continued to improve their knowledge of contraceptives during their second summer of participation. Their knowledge of substance abuse was positively affected as well.

COSTS

The social costs of enrolling a youth in the 15-month STEP program are derived from analysis of a synthetic (hypothetical) cohort. The estimated 15-month net cost per STEP enrollee, over the cost of providing work experience alone for two summers, was $1,586.[2] Three of the sites--Boston, Fresno and Portland--had total net costs per enrollee of between $1,200 and $1,500. San Diego and Seattle had per enrollee net costs of between $1,600 and $1,700 (despite the lack of a school-year support program in San Diego). Our best net cost estimate of operating a mature STEP intervention that included an active school-year support program is approximately $1,600 per enrollee (for sites enrolling approximately 150 new participants per year).

These costs are the social costs and do not necessarily reflect the actual site costs of running the program. For example, they omit transfer payments--wages paid to enrollees, (and associated administrative costs) which are roughly $1,400 over the 15-month period. Further, they include the costs of contributed space and computers, for which almost no site had to pay.

IMPLICATIONS

The STEP experience and research to date provide some important insight into the need, viability and usefulness of summer-based strategies to improve basic skills and keep youth in school.

STEP's corroboration and extension of previous research on summer learning loss indicate the need. The size and consistency of losses in reading and math experienced by control group youth in both cohorts is startling. The magnitude of these summer learning losses seems disproportionate to the brief period of time that elapsed while youth were working on SYETP jobs.

---

[2] The average cost of providing only one summer of remediation and LSO instruction is $717 per enrollee.

Though there is still a great deal to be learned about the learning loss phenomenon--whether or not it is sensitive to the type of test administered, and if and how summer losses can be retrieved during successive school years--it calls for a re-assessment of the educational importance of the summer period.

Such a reassessment would focus on the summer's emptiness of content vis a vis basic skills, rather than on its brevity as an opportunity for intervention. STEP, in fact, lasted only seven weeks--with remediation offered for about five half-days a week--and yet it was sufficient to stem substantial losses in reading and produce actual gains in math.

The nationwide summer jobs program provides one opportunity to test alternatives. Starting this summer, localities are required to assess the basic skills of the 650,000 youth enrolled in the program and to provide remedial assistance to those with weak basic skills. A survey by the National Job Training Partnership found that the vast majority of localities are organizing or planning to organize a remediation component along with their summer work experience programs in 1987.

The STEP experience in remediation--particularly the differences in approach and test results between the 1985 and 1986 summers--should offer some useful insight into the level of resources and educational structures necessary to produce short-term test results that hold promise for long-term improvements in basic skills.

STEP's operational experience and test results seem to confirm both the feasibility and importance of extended educational programming for high-risk students. STEP differs from simple school-year extension, however, by integrating into the summer's academic instruction an opportunity for low income youth to work and earn a salary, and to discuss and learn about key life options and their implications. These additional elements may be crucial in providing the economic incentive and practical knowledge necessary for continued participation in regular schooling.

Finally, the STEP experience to date demonstrates the feasibility of public education and employment/training institutions working together to provide innovative, multi-dimensional and effective programming for high risk youth.

REFERENCES

Amemiya, T. "Qualitative Response Models," Journal of Economic Literature, 1981, Vol. 19:1438-56.

Berlin, Gordon, Andrew Sum and Robert Taggart, "Cutting Through." Unpublished Manuscript, 1986.

Betsey, Charles L., Robinson G. Hollister, Jr. and Mary R. Papageorgiou, Eds., Youth Employment and Training Programs: The YEDPA Years. Committee on Youth Employment Programs, Commission on Behavioral and Social Sciences and Education, National Research Council. National Academy Press, Washington, D.C.: 1985.

Branch, Alvia Y., Julita Milliner and Jon Bumbaugh, Summer Training and Education Program (STEP): Report on the 1985 Summer Experience. Public/Private Ventures, Philadelphia, Pennsylvania: April 1986.

Branch, Alvia Y., Julita Milliner, Susan Phillipson Bloom and Jon Bumbaugh, Summer Training and Education Program: Report on the Pilot Experience. Public/Private Ventures, Philadelphia, Pennsylvania: March 1985.

Card, Josefina J. and Lauress L. Wise, "Teenage Mothers and Teenage Fathers: The Impact of Early Childbearing on the Parents' Personal and Professional Lives," Family Planning Perspective, 1978, Vol. 10:199-205.

Coombs, J. and W.W. Cooley, "Dropouts: In High School and After High School," American Educational Research Journal, 1968, Vol. 5:343-358.

David, Jane and Sol H. Pelavin, "Secondary Analysis: In Compensatory Education Programs," New Directions for Program Evaluation, 1978, Vol. 4:31-44.

Dryfoos, J.G., Prevention Strategies: A Progress Report. Report to the Rockefeller Foundation: 1985.

Furstenberg, F.F., Jr., Unplanned Parenthood: The Social Consequences of Teenage Childbearing. Free Press, New York: 1976.

Hammond, Pierce A. III and Joy A. Frechtling, "Twelve, Nine and Three Month Achievement Gains of Low and Average Achieving Elementary School Students." Paper presented at the annual meeting of the American Educational Research Association, San Francisco, California: April 1979.

Hayes, Cheryl D., Ed., *Risking the Future: Adolescent Sexuality, Pregnancy and Child Bearing*. Report of the Panel on Adolescent Pregnancy and Childbearing, National Research Council. National Academy Press, Washington, D.C.: 1987.

Hayes, Donald P. and Judith Grether, "The School Year and Vacations: When Do Students Learn?" Revision of a paper presented at the Eastern Sociological Association Convention, New York City, April, 1969.

Heyns, Barbara, *Summer Learning and the Effects of Schooling*. Academic Press, New York: 1978.

Heyns, Barbara, "Schooling and Cognitive Development: Is There a Season for Learning?" Unpublished Manuscript, 1986.

Institute for Educational Leadership, Inc., *School Dropouts: Everybody's Problem*. Report of the Institute for Educational Leadership Conference, Washington, D.C.: 1986.

Kapsis, Robert and William Protash, *Summer Motivation and Retention*. City University of New York: 1978.

Kirsch, Irwin S. and Ann Jungeblut, *Literacy: Profiles of America's Young Adults*. Report No. 16-PL-02, National Assessment of Educational Progress, Princeton, New Jersey: 1986.

Levin, Henry M., *The Educationally Disadvantaged: A National Crisis*. Public/Private Ventures, Philadelphia, Pennsylvania: July 1985.

Malizio, Andrew G. and Douglas R. Whitney, *Educational Credentials in Employment: A Nationwide Survey*. Paper presented at the Lifelong Learning Conference, College Park, Maryland: February 16-17, 1984.

McClendon, M.J., "The Occupational Status Attainment Processes of Males and Females," *American Sociological Review*, 1976, Vol. 41:52-64.

National Academy of Sciences, *High Schools and the Changing Workplace*, Report of the Panel on Secondary School Education for the Changing Workplace. National Academy Press, Washington, D.C.: 1984.

National Center for Education Statistics, U.S. Department of Education, *The Condition of Education*, 1985 Edition. U.S. Government Printing Office, Washington, D.C.: 1985.

National Job Training Partnership, Inc., *Education in Local JTPA Summer Youth Programs: A Fuller National Picture*. Washington, D.C.: August 1986.

Pelavin, Sol H. and Jane David, *Evaluating Long-Term Achievement: An Analysis of Longitudinal Data from Compensatory Education Programs*. Research Report, SRI Project 4537. Stanford Research Institute, International, Menlo Park, California: March 1977.

Presser, H.B., "Social Consequences of Teenage Childbearing," in C. Chilman (ed.), *Teenage Childbearing: Recent Research on the Determinants and Consequences*. U.S. Government Printing Office, Washington, D.C.: 1978.

Quint, Janet C. and James A. Riccio, *The Challenge of Serving Pregnant and Parenting Teens: Lessons from Project Redirection*. Manpower Demonstration Research Corporation, New York: April 1985.

Sum, Andrew, et al., *Educational Attainment, Academic Ability, and the Employability and Earnings of Young Persons: Implications for the Planning and Design of JTPA Youth Programs*. Prepared for a Symposium on Serving Disadvantaged School Dropouts. Brandeis University: December 1983.

Taggart, Robert, *A Fisherman's Guide: An Assessment of Training and Remediation Strategies.* W.E. Upjohn Institute, Kalamazoo, Michigan: 1981.

U.S. Department of Commerce, Bureau of the Census, *Lifetime Earnings Estimates for Men and Women in the United States*, Current Population Reports, Series p. 60, No. 139. U.S. Government Printing Office, Washington, D.C.: 1983.

Walker, Gary, *JTPA: Early Implementation*. Grinker, Walker and Associates, New York: 1985.

APPENDIX A

PROGRAM PARTICIPATION ANALYSIS

APPENDIX A

PROGRAM PARTICIPATION ANALYSIS

Before analyzing program impacts, we examined the determinants of program participation. This analysis considers what characteristics predict how much remediation and work the youth receive. Factors found to be good predictors of program participation may enable us to adjust the program's target population in order to increase the number of youth who complete the program.

For treatment youth, program hours were divided between classroom hours (including remediation and life skills classes) and work hours. As discussed in Chapter IV, Cohort II treatment youth attended an average of 100 classroom hours. Multivariate analysis was used to investigate the influence of a number of variables on classroom hours attended by treatment youth. Results of this analysis indicate that there were no significant differences in classroom attendance by sex or age. Asian youth attended significantly more hours than other race/ethnic groups; and youth with better basic reading skills at baseline and higher educational expectations tended to have better attendance. The best predictor of classroom attendance, however, was the youth's attendance pattern during the preceding school year. Youth who reported being absent fewer than 10 days during the 1985-86 school year had better remediation/life skills attendance during STEP, than youth with high absenteeism during the school year. Thus it appears that the most motivated youth (as indicated by school-year attendance) get more exposure to the STEP treatment.

The second indicator of program participation was total hours attended. Total hours consist only of work hours for control youth, but include both classroom and work hours for treatment youth. As discussed in Chapter IV, Cohort II control youth averaged 206 total hours of participation while treatment youth in Cohort II participated in an average of 207 total program hours. The multivariate analysis of total program hours of participation indicated that treatment status, sex, age or race/ethnicity did not significantly influence total program hours. Youth who indicated that they expected to continue their education beyond high school and those with English language problems had better attendance. But, as was true for remediation attendance, the best predictor of program hours was regular school attendance. Youth who had more than 10 absences during the preceding school year averaged 14 fewer program hours than youth with good school-year attendance. Thus youth who are motivated to attend school also appear to be more motivated during the summer.

Table A.1

DETERMINANTS OF PROGRAM HOURS

Classroom Hours		Total Hours	
Explanatory Variable	Coefficient	Explanatory Variable	Coefficient
Constant	-59.34***	Constant	120.21***
Sex (male = 1)	-1.42	Treatment	3.43
Black[a]	4.68	Sex (male = 1)	-4.29
Hispanic[a]	4.68	Black[a]	2.19
Asian[a]	9.59**	Hispanic[a]	6.74
Family Size	0.60	Asian[a]	4.77
Birth Order	-0.30	Family Size	1.62**
Grade Previous Year	-1.06	Birth Order	-0.42
Repeated a Grade	-0.01	Grade Previous Year	2.01
School Absences	-10.91***	Repeated a Grade	-0.58
Expects Post-high School Education	5.21**	School Absences	-14.16***
English Language Problem	5.50	Expects Post-high School Education	8.23**
Baseline Reading	0.04**	English Language Problem	9.74**
Boston[b]	27.15***	Baseline Reading	0.08***
Fresno[b]	17.24***	Boston[b]	-28.28***
Portland[b]	-5.35*	Fresno[b]	31.57***
San Diego[2]	9.03***	Portland[b]	-41.70***
		San Diego[2]	13.47***
Sample Size	637		1,268

[a] Race/ethnicity subgroup impacts are relative to the excluded category, whites and others.
[b] Site impacts are relative to the excluded category, Seattle.

***Indicates that the impact statistically differs from zero at the 0.01 level.
**Indicates that the impact statistically differs from zero at the 0.05 level.
*Indicates that the impact statistically differs from zero at the 0.10 level.

APPENDIX B

SUMMER WORK ENHANCEMENT PROGRAM (SWEP)

APPENDIX B

SUMMER WORK ENHANCEMENT PROGRAM (SWEP)

During the first summer of the demonstration, STEP was successful in stemming learning losses, but did not produce actual learning gains among treatment youth as indicated by their scores on the Metropolitan Achievement Test (MAT). (In the second summer, however, STEP treatment youth did experience small learning gains in math.) Other remedial programs, in particular the Comprehensive Competencies Program (CCP), have reported learning gains as measured by the Test of Adult Basic Education (TABE). A question arose, therefore, about whether the lack of learning gains in STEP could be an artifact of the testing instrument.

The Summer Work Enhancement Program (SWEP) operated by the New York City Private Industry Council in the summer of 1986 provided an opportunity to test a group of students with both the MAT and the TABE. SWEP youth received a mix of remediation, life skills and work experience similar to STEP. The SWEP remediation curriculum was the CCP developed by Robert Taggart; the Life Skills and Opportunities curriculum was developed by P/PV. As shown in Table B.1, the 21 SWEP students were like STEP youth, primarily from minority groups, but tended to be a little older. Their baseline test scores (Table B.2) indicate that they were also educationally disadvantaged. Both STEP and SWEP students were performing at approximately the sixth grade level in reading and between the sixth and seventh grade level in math when they entered the programs.

Participants in SWEP took both the MAT and the TABE. On the first day of testing, youth completed the STEP baseline questionnaire and then the MAT intermediate level reading and math tests. This sequence is identical to that used for testing at STEP sites. Thus, SWEP students took the MAT under conditions as similar as possible to those in STEP sites. On the second day of the program, youth completed the TABE in accordance with usual CCP testing procedures. At the end of the summer program, youth completed the endline questionnaire used by STEP, the MAT and the TABE.

SWEP administrators forwarded all test and questionnaire data to P/PV for analysis. Table B.2 presents mean pre-test scaled scores on both the MAT and the TABE, and the corresponding grade equivalents for the sample of SWEP youth with complete data on

Table B.1

SELECTED CHARACTERISTICS OF STEP COHORTS I AND II AND SWEP PARTICIPANTS

	STEP		
	Cohort I	Cohort II	SWEP
Sex			
Male	53%	49%	43%
Female	47	51	57
Race/Ethnicity			
Black	59%	48%	58%
Hispanic	16	18	37
Asian	14	19	-
White & Other	11	15	5
Age			
14 or Younger	50%	55%	5%
15	50	45	48
16 or Older	-	-	47
Grade			
8th or Below	50%	60%	5%
9th	50	35	71
10th or Above	-	5	24
Sample Size	1,495	1,268	21

Table B.2

MEAN TEST SCORE CHANGES FOR STEP COHORTS I AND II AND SWEP PARTICIPANTS

	Reading		Math	
	Pre-test	Change	Pre-test	Change
COHORT I				
Treatments				
MAT Scaled Score	713.31	-8.55***	690.21	-0.24
Grade Equivalent	5.8	-0.4	6.5	0.0
Controls				
MAT Scaled Score	718.95	-14.32***	695.96	-16.47***
Grade Equivalent	5.9	-0.7	6.6	-0.4
COHORT II				
Treatments				
MAT Scaled Score	721.19	-6.67***	700.51	7.03***
Grade Equivalent	6.0	-0.3	6.8	0.2
Controls				
MAT Scaled Score	723.69	-19.50***	706.08	-17.73***
Grade Equivalent	6.2	-1.0	6.9	-0.5
SWEP PARTICIPANTS				
MAT Scaled Score	723.90	-7.28	675.52	-1.33
Grade Equivalent	6.2	-0.4	6.1	0.0
TABE Scaled Score	446.24	22.14***	454.14	10.72
Grade Equivalent	5.8	0.8	6.7	0.3

***Indicates that the impact statistically differs from zero at the 0.01 level.

both tests.[1] For comparison purposes, first summer data for STEP Cohorts I and II are also presented.

As noted above, SWEP and STEP students tested at roughly comparable levels in both reading and math at the beginning of the summer. The mean changes on each test from the pre-test to the post-test are also presented in Table B.2. As measured by the MAT, SWEP students lost 0.4 of a grade equivalent in reading, nearly identical to the reading losses of STEP treatment youth in both cohorts. As measured on the MAT, the test score level for both SWEP students and STEP Cohort I treatments did not change over the summer, while STEP Cohort II treatment youth gained 0.2 of a grade in math.

When the change in skill levels for SWEP students was measured by the TABE, however, the results are quite different. On the TABE, SWEP youth, who had shown a loss of 0.4 on the MAT, gained an average of 0.8 of a grade in reading.

P/PV consulted several testing experts for an interpretation of these findings and an exploration of their implications for our work.

Clearly, test results can vary. Our consultants noted that the SWEP students' pattern of gains on one test and losses on another is common in remedial programs. First, tests can measure quite different skills; it is reputed, for example, that the MAT is an academic test while the TABE measures more practical skills. In addition, the match between the content of the test items and the content of the instructional program plus the relative difficulty of the two tests contribute to differential results. Thus, an evaluation of a program's effectiveness should concentrate on the performance of treatment youth relative to that of controls, rather than on the absolute performance (gains or losses) of treatments.

Nevertheless, the SWEP results leave open the possibility that the large losses STEP control youth experienced on the MAT might have been scored as "no change" over the summer on the TABE. This does not suggest, however, that the control groups losses measured by the MAT are artifactual and do not represent "real" learning losses. It is plausible that the MAT and the TABE are measuring the amount of summer loss for different skill areas.

---

[1] Because the MAT and TABE scaled scores are in different metrics, they are not comparable. That is, a change of 10 points on the scaled scores of one test is not equivalent to a change of 10 points on the other test. Scores are presented in terms of grade equivalents to provide a means of comparing the results from the two tests.

APPENDIX C

TABLES REFERENCED IN TEXT

Chapter V: Participant Characteristics
Chapter VI: First Summer Impacts
Chapter VII: School-Year and Second Summer Impacts

Table C.1

DEMOGRAPHIC AND EDUCATIONAL ABILITY CHARACTERISTICS OF THE AGGREGATE SAMPLE

Characteristic	All Sites	Boston	Fresno	Portland	San Diego	Seattle
Age						
14 Years Old at Entry	52%	57%	54%	54%	53%	45%
15 Years Old at Entry	48	43	46	46	47	55
Sex						
Male	50%	54%	53%	51%	47%	48%
Female	50	46	47	49	53	52
Race/Ethnicity						
Asian	17%	4%	7%	13%	32%	29%
Black	53	64	40	58	44	59
Hispanic	17	17	44	3	18	2
White	9	12	6	20	3	6
Other	4	3	3	6	3	4
Educational Ability						
Mean MAT Reading Score at Entry (grade equivalent)	721 (6.0)	716 (5.7)	721 (6.0)	724 (6.2)	730 (6.5)	713 (5.6)
Mean MAT Math Score at Entry (grade equivalent)	699 (6.6)	684 (6.3)	703 (6.7)	696 (6.6)	734 (7.8)	681 (6.2)
Sample Size	2,431	504	486	473	483	485

Table C.2

ADDITIONAL DEMOGRAPHIC AND ECONOMIC CHARACTERISTICS OF THE AGGREGATE SAMPLE

Characteristic	All Sites	Boston	Fresno	Portland	San Diego	Seattle
Percent From Female-Headed Households	50%	62%	45%	49%	47%	49%
Mean Household Size	5.21	5.04	5.31	4.87	5.76	5.08
Number Married	9	0	4	1	3	1
Number With Children	22	1	5	7	3	6
Percent Ever Worked	32%	35%	33%	39%	25%	27%
Sample Size	2,431	504	486	473	483	485

Table C.3

MEAN MAT READING SCORES AT ENTRY FOR THE AGGREGATE SAMPLE BY AGE, RACE/ETHNICITY AND SEX

Characteristic	All Sites	Boston	Fresno	Portland	San Diego	Seattle
Age						
14 Years Old	720	723	717	723	729	710
15 Years Old	721	708	727	725	732	716
Race/Ethnicity						
Asian	719	722	725	727	731	699
Black	720	718	713	721	728	719
Hispanic	724	709	725	721	734	740
White	725	719	744	727	731	712
Others	725	716	724	733	730	719
Sex						
Male	715	709	723	717	724	704
Female	727	722	720	731	736	723
Sample Size	2,431	504	486	473	483	485

Table C.4

MEAN MAT MATH SCORES AT ENTRY FOR THE AGGREGATE SAMPLE BY AGE, RACE/ETHNICITY AND SEX

Characteristic	All Sites	Boston	Fresno	Portland	San Diego	Seattle
Age						
14 Years Old	697	692	692	695	730	675
15 Years Old	702	673	715	696	739	686
Race/Ethnicity						
Asian	748	742	740	756	772	719
Black	683	681	689	682	710	663
Hispanic	706	684	706	687	729	713
White	694	680	720	699	705	673
Other	692	679	703	685	729	676
Sex						
Male	698	680	706	696	732	678
Female	700	686	698	695	736	685
Sample Size	2,431	504	486	473	483	485

Table C.5

SCHOOL PARTICIPATION AND EDUCATIONAL EXPECTATIONS OF THE AGGREGATE SAMPLE

Characteristic	All Sites	Boston	Fresno	Portland	San Diego	Seattle
Grade in School						
8th grade or below	56%	85%	43%	44%	52%	51%
9th grade	38	14	51	51	41	37
10th grade	6	1	6	5	6	12
11th grade	<1	0	0	0	1	<1
Grade Retention History						
Percent Ever Repeating a Grade	30%	45%	32%	22%	22%	26%
Absenteeism						
Percent Absent 1-10 Days in SY 85-86	73%	74%	63%	72%	81%	74%
Percent Absent 11-20 Days in SY 85-86	18	17	23	21	13	16
Percent Absent More Than 20 Days in SY 85-86	9	9	14	7	6	10
Educational Expectations						
Percent Planning to Pursue Post-High School Education	72%	62%	74%	75%	76%	74%
Sample Size	2,431	504	486	473	483	485

Table C.6

FERTILITY-RELATED CHARACTERISTICS OF THE AGGREGATE SAMPLE

Characteristic (Sample Size)	All Sites	Boston	Fresno	Portland	San Diego	Seattle
Percent With Prior Sexual Experience (2,247)	45%	53%	40%	52%	36%	43%
Percent of the Sexually Active Sample Who Were Sexually Active in the Past Two Months (975)	54%	56%	52%	55%	52%	55%
Percent of Recently Active Sample Never Using Contraception During the Past Two Months (515)	38%	43%	42%	32%	47%	28%
Mean Age at Onset of Sexual Activity for the Sexually Active Sample (1,019)	11.92	11.95	12.24	11.75	11.80	11.91
Number of Females Ever Pregnant (341)	34	5	3	12	3	11
Number of Males Ever Fathered a Child or Responsible for a Pregnancy (670)	32	5	4	8	5	10

Table C.6 - (cont'd)

FERTILITY-RELATED CHARACTERISTICS OF THE AGGREGATE SAMPLE

Characteristic (Sample Size)	All Sites	Boston	Fresno	Portland	San Diego	Seattle
Mean Number Correct on Birth Control Availability Questions (8-item scale)	3.07	3.11	2.98	3.29	3.06	2.95
Mean Number Correct on General Birth Control Knowledge Questions (12-item scale)	4.97	4.94	5.06	5.12	4.88	4.85
Percent Believing That Both Partners Should Be Responsible for Birth Control (2,404)	73%	65%	72%	78%	75%	75%
Percent Who have Taken a Prior Sex Education Course (2,267)	65%	54%	57%	71%	84%	60%

Table C.7

MEAN MAT READING SCORES AT ENTRY FOR COHORT II BY AGE, RACE/ETHNICITY AND SEX

Characteristic	All Sites	Boston	Fresno	Portland	San Diego	Seattle
Age						
14 Years Old	720	711***	718	725	730	717
15 Years Old	725	695	741	728	732	731
Race/Ethnicity						
Asian	721	709	725	729	730	709
Black	722	707	719	725	730	731
Hispanic	722	696	728	716	734	752
White	727	711	753	722	728	733
Other	725	699	745	734	729	715
Sex						
Male	716***	692***	729	719**	722***	717*
Female	728	715	726	733	739	731
Sample Size	1,268	262	256	246	258	246

***Treatment-control differences are significantly different from zero at the 0.01 level.

**Treatment-control differences are significantly different from zero at the 0.05 level.

*Treatment-control differences are significantly different from zero at the 0.10 level.

Table C.8

MEAN MAT MATH SCORES AT ENTRY FOR COHORT II BY AGE, RACE/ETHNICITY AND SEX

Characteristic	All Sites	Boston	Fresno	Portland	San Diego	Seattle
Age						
14 Years Old	701	688	700	700	730	689
15 Years Old	706	666	723	702	736	699
Race/Ethnicity						
Asian	750***	724	740	754	769	735
Black	688	677	689	690	713	678
Hispanic	703	673	712	665	720	702
White	698	679	718	702	693	684
Other	695	663	723	686	731	685
Sex						
Male	701	672	712	702	728	688
Female	705	684	707	699	737	700
Sample Size	1,268	262	256	246	258	246

***Treatment-control differences are significantly different from zero at the 0.01 level.

Table C.9

DEMOGRAPHIC CHARACTERISTICS AND EDUCATIONAL ABILITY LEVEL OF RETURNEES IN SAN DIEGO

	Controls		Treatments	
	Non-returnees	Returnees	Non-returnees	Returnees
Age				
14 Years Old at Entry	58%	48%	42%	60%
15 Years Old at Entry	42	52	58	40
Sex				
Male	46%	45%	53%	51%
Female	54	55	47	49
Race/Ethnicity				
Asian	27%	33%	19%	35%
Black	40	49	61	45
Hispanic	27	15	17	16
White & Other	6	3	3	4
Educational Ability				
Mean MAT Reading Score at Entry (grade equivalent)	717 (5.8)	729 (6.4)	726 (6.3)	729 (6.4)
Mean MAT Math Score at Entry (grade equivalent)	733 (7.8)	730 (7.6)	714 (7.1)	731 (7.7)
Sample Size	52	102	79	77

Table C.10

ADDITIONAL DEMOGRAPHIC AND ECONOMIC CHARACTERISTICS OF RETURNEES IN SAN DIEGO

	Controls		Treatments	
	Non-returnees	Returnees	Non-returnees	Returnees
Percent With Limited English Proficiency	19%	16%	11%	17%
Percent From Female-Headed Households	51%	45%	47%	47%
Mean Household Size	5.77	5.73	4.93	5.68
Percent Ever Worked	38%	31%	29%	28%
Sample Size	52	102	79	77

Table C.11

SCHOOL PARTICIPATION CHARACTERISTICS AND EDUCATIONAL EXPECTATIONS OF RETURNEES IN SAN DIEGO

	Controls		Treatments	
	Non-returnees	Returnees	Non-returnees	Returnees
Grade in School				
8th Grade or Below	48%	50%	48%	51%
9th Grade	48	45	37	43
10th Grade	4	5	15	5
Grade Retention History				
Percent Ever Repeating a Grade	15%	25%	30%	20%
Absenteeism				
Percent Absent 1-10 Days in SY 84-85	74%	81%	83%	78%
Percent Absent 11-20 Days in SY 84-85	16	12	9	16
Percent Absent More Than 20 Days in SY 84-85	10	7	8	6
Educational Expectations				
Percent Planning to Pursue Post-High School Education	61%	77%	86%	80%
Sample Size	52	102	79	77

Table C.12

FERTILITY-RELATED CHARACTERISTICS OF RETURNEES IN SAN DIEGO

	Controls		Treatments	
Characteristic	Non-returnees	Returnees	Non-returnees	Returnees
Percent With Prior Sexual Experience	34%	31%	61%	34%
Percent of the Sexually Active Sample Who Were Sexually Active in the Past Two Months	75%	55%	63%	42%
Percent of the Recently Active Sample Never Using Contraception During Past Two Months	67%	40%	38%	50%
Mean Age at Onset of Sexual Activity for the Sexually Active Sample	11.92	11.17	12.10	11.41
Mean Number Correct on Birth Control Availability Questions (8-item scale)	2.73	3.10	4.57	2.76
Mean Number Correct on General Birth Control Knowledge Questions (12-item scale)	4.33	5.12**	4.57	4.31**
Percent Who Have Taken a Prior Sex Education Course	73%	79%	88%	73%

**The treatment-control difference is significantly different from zero at the 0.05 level.

Table C.13

PRE-TEST TO POST-TEST CHANGES IN MEAN READING AND MATH SCORES BY SEX AND RACE/ETHNICITY FOR AGGREGATE SAMPLE

Group (Sample Size)	Reading Scores		Math Scores	
	Change from Pre- to Post-test	Treatment-Control Difference	Change from Pre- to Post-test	Treatment-Control Differences
Total Sample				
Treatment (1,219)	-6.82***		3.80**	
Control (1,212)	-16.87***	10.05***	-16.79***	20.59***
Males				
Treatment (592)	-10.29***		-2.16	
Control (619)	-21.59***	11.30***	-27.25***	25.09***
Females				
Treatment (627)	-3.54**		9.42***	
Control (593)	-11.94***	8.40***	-5.88**	15.30***
Asians				
Treatment (211)	7.53***		15.91***	
Control (202)	-2.43	9.96***	-3.71	19.62***
Blacks				
Treatment (663)	-10.85***		2.45	
Control (626)	-19.80***	8.95***	-19.09***	21.54***
Hispanics				
Treatment (204)	-5.90*		-2.58	
Control (203)	-20.23***	14.33***	-25.04***	22.46***
Whites & Others				
Treatment (141)	-10.65**		1.24	
Control (181)	-19.07***	8.42	-14.20**	15.44**

***Difference is significantly different from zero at the 0.01 level.
**Difference is significantly different from zero at the 0.05 level.
*Difference is significantly different from zero at the 0.10 level.

Table C.14

PRE-TEST TO POST-TEST CHANGES IN MEAN READING AND MATH SCORES BY SITE FOR AGGREGATE SAMPLE

Group (Sample Size)	Reading Scores: Change from Pre- to Post-test	Reading Scores: Treatment-Control Difference	Math Scores: Change from Pre- to Post-test	Math Scores: Treatment-Control Differences
Total Sample				
Treatment (1,219)	-6.82***		3.80**	
Control (1,212)	-16.87***	10.05***	-16.79***	20.59***
Boston				
Treatment (252)	0.32		8.91**	
Control (252)	-18.43***	18.75***	-23.65***	32.56***
Fresno				
Treatment (248)	-9.96***		-1.66	
Control (238)	-17.03***	7.07*	-25.35***	23.69***
Portland				
Treatment (241)	-9.97***		2.48	
Control (232)	-13.13***	3.16	-8.17**	10.65*
San Diego				
Treatment (235)	-4.74		3.67	
Control (248)	-11.06***	6.32	-4.63	8.30
Seattle				
Treatment (243)	-9.89***		5.50	
Control (242)	-24.61***	14.72***	-21.97***	27.47***

***Difference is significantly different from zero at the 0.01 level.
**Difference is significantly different from zero at the 0.05 level.
*Difference is significantly different from zero at the 0.10 level.

Table C.15

BASELINE TO ENDLINE CHANGES IN MEAN KNOWLEDGE OF CONTRACEPTION AND AVAILABILITY OF CONTRACEPTION FOR COHORT II OVERALL AND BY SEX, RACE/ETHNICITY, SITE

	Contraceptive Knowledge			Contraceptive Availability		
	Baseline	Endline	Change	Baseline	Endline	Change
Treatments	5.08	7.75	2.67***	3.03	5.16	2.13***
Controls	5.04	5.12	0.08	3.15	3.45	0.30***
Difference			2.59***			1.83***
Males						
Treatments	5.20	7.77	2.57***	2.99	5.13	2.14***
Controls	5.02	5.05	0.03	3.04	3.16	0.12
Difference			2.54***			2.02***
Females						
Treatments	4.97	7.72	2.75***	3.07	5.17	2.10***
Controls	5.07	5.19	0.12	3.26	3.76	0.50***
Difference			2.63			1.60***
Asians						
Treatments	3.49	7.52	4.03***	2.23	5.50	3.27***
Controls	3.42	3.47	0.05	2.06	2.29	0.23
Difference			3.98***			3.04***
Blacks						
Treatments	5.51	7.88	2.37***	3.22	5.03	1.81***
Controls	5.45	5.43	-0.02	3.41	3.62	0.21*
Difference			2.39***			1.60***
Hispanics						
Treatments	5.19	7.56	2.37***	3.21	5.28	2.07***
Controls	5.28	5.49	0.21	3.36	3.90	0.54***
Difference			2.16***			1.53***

Table C.15 - (cont'd)

BASELINE TO ENDLINE CHANGES IN MEAN KNOWLEDGE OF CONTRACEPTION AND AVAILABILITY OF CONTRACEPTION FOR COHORT II OVERALL AND BY SEX, RACE/ETHNICITY, SITE

	Contraceptive Knowledge			Contraceptive Availability		
	Baseline	Endline	Change	Baseline	Endline	Change
Whites & Others						
Treatments	5.56	7.85	2.29***	3.24	4.98	1.74***
Controls	5.36	5.58	0.22	3.34	3.75	0.41***
Difference			2.07***			1.33***
Boston						
Treatments	4.47	6.10	1.63***	2.77	4.12	1.35***
Controls	5.00	5.27	0.27	3.20	3.80	0.60***
Difference			1.36***			0.75***
Fresno						
Treatments	5.30	7.88	2.58***	2.96	4.99	2.03***
Controls	5.17	5.39	0.22	3.17	3.69	0.52***
Difference			2.36***			1.51***
Portland						
Treatments	5.33	8.12	2.79***	3.22	5.29	2.07***
Controls	5.04	5.12	0.08	3.25	3.26	0.01
Difference			2.71***			2.06***
San Diego						
Treatments	4.95	9.65	4.70***	3.05	6.53	3.48***
Controls	5.20	5.04	-0.16	3.17	3.33	0.16
Difference			4.86***			3.32***
Seattle						
Treatments	5.36	6.96	1.60***	3.18	4.83	1.65***
Controls	4.82	4.76	-0.06	2.96	3.18	0.22
Difference			1.66***			1.43***

***Indicates that change statistically differs from zero at the 0.01 level.

Table C.16

BASELINE TO ENDLINE CHANGES IN MEAN KNOWLEDGE OF CONTRACEPTION AND AVAILABILITY OF CONTRACEPTION FOR AGGREGATE SAMPLE OVERALL BY SEX, RACE/ETHNICITY, SITE

	Contraceptive Knowledge			Contraceptive Availability		
	Baseline	Endline	Change	Baseline	Endline	Change
Treatments	4.95	7.55	2.60***	3.04	5.06	2.02***
Controls	5.01	5.09	0.08	3.13	3.39	0.26***
Difference			2.52***			1.76***
Males						
Treatments	4.95	7.45	2.50***	2.92	4.91	1.99***
Controls	5.06	5.12	0.06	3.08	3.21	0.13*
Difference			2.44***			1.86***
Females						
Treatments	4.94	7.64	2.70***	3.16	5.20	2.04***
Controls	4.95	5.05	0.10	3.18	3.57	0.39***
Difference			2.60***			1.65***
Asians						
Treatments	3.54	6.78	3.24***	2.22	4.88	2.66***
Controls	3.56	3.58	0.02	2.10	2.27	0.17
Difference			3.22***			2.49***
Blacks						
Treatments	5.33	7.66	2.33***	3.24	5.01	1.77***
Controls	5.29	5.43	0.14*	3.31	3.52	0.21***
Difference			2.19***			1.56***
Hispanics						
Treatments	4.95	7.82	2.87***	3.16	5.46	2.30***
Controls	5.28	5.30	0.02	3.26	3.72	0.46***
Difference			2.85***			1.84***

Table C.16 - (cont'd)

BASELINE TO ENDLINE CHANGES IN MEAN KNOWLEDGE OF CONTRACEPTION AND AVAILABILITY OF CONTRACEPTION FOR AGGREGATE SAMPLE OVERALL BY SEX, RACE/ETHNICITY, SITE

	Contraceptive Knowledge			Contraceptive Availability		
	Baseline	Endline	Change	Baseline	Endline	Change
Whites & Others						
Treatments	5.28	7.76	2.48***	3.17	4.99	1.82***
Controls	5.29	5.31	0.02	3.46	3.78	0.32***
Difference			2.46***			1.50***
Boston						
Treatments	4.77	6.61	1.84***	3.01	4.48	1.47***
Controls	5.04	5.24	0.20	3.18	3.70	0.52***
Difference			1.64***			0.95***
Fresno						
Treatments	5.01	8.15	3.14	2.86	5.25	2.39***
Controls	5.18	5.28	0.10	3.12	3.45	0.33***
Difference			3.04***			2.06***
Portland						
Treatments	5.17	7.64	2.47***	3.23	5.00	1.77***
Controls	5.02	5.20	0.18	3.34	3.40	0.06
Difference			2.29***			1.71***
San Diego						
Treatments	4.71	8.68	3.97***	3.01	6.03	3.02***
Controls	5.08	5.04	-0.04	3.13	3.29	0.16
Difference			4.01***			2.86***
Seattle						
Treatments	5.10	6.66	1.56***	3.11	4.55	1.44***
Controls	4.72	4.69	-0.02	2.86	3.10	0.24**
Difference			1.58***			1.20***

***Indicates that change statistically differs from zero at the 0.01 level.
**Indicates that change statistically differs from zero at the 0.05 level.
*Indicates that change statistically differs from zero at the 0.10 level.

Table C.17

THE ESTIMATED IMPACTS OF STEP ON THE PERCENTAGE OF THE 1985-86 SCHOOL YEAR ABSENT BY SITE FOR COHORT I

Sample Site	Estimated Impact	Control Group Mean	Sample Size
Boston	-2.89%	16.85%	244
Fresno	0.71	13.96	251
Portland	-1.14	13.48	216
San Diego	na	na	na
Seattle	-0.45	7.35	217

na = San Diego absenteeism data not available.

***Indicates that impact statistically differs from zero at the 0.01 level.
**Indicates that impact statistically differs from zero at the 0.05 level.
*Indicates that change statistically differs from zero at the 0.10 level.

Table C.18

THE ESTIMATED IMPACT OF STEP ON THE PERCENTAGE OF GRADUATION CREDITS EARNED DURING THE INTERVENING SCHOOL YEAR BY SITE FOR COHORT I

Site	Estimated Impact (Percentage Points)	Sample Size
Boston	0.81	216
Fresno	-1.13	228
Portland	-1.81	188
San Diego	-0.58	144
Seattle	3.02	184

***Indicates that impact statistically differs from zero at the 0.01 level.

**Indicates that impact statistically differs from zero at the 0.05 level.

*Indicates that impact statistically differs from zero at the 0.10 level.

Table C.19

THE ESTIMATED IMPACT ON THE PROBABILITY OF PROMOTION DURING THE INTERVENING SCHOOl YEAR BY SITE FOR COHORT I

Site	Estimated Impact (Percentage Points)	Percent of Control Group Promoted	Sample Size
Boston	9.22	64.97	244
Fresno	6.10	68.92	254
Portland	7.02	80.15	222
San Diego	1.17	92.65	217
Seattle	-3.50	76.39	230

***Indicates that impact statistically differs from zero at the 0.01 level.
**Indicates that impact statistically differs from zero at the 0.05 level.
*Indicates that impact statistically differs from zero at the 0.10 level.

Table C.20

THE ESTIMATED IMPACTS OF STEP ON NORMAL CURVE EQUIVALENTS (NCEs) OF TEST SCORES BY SEX, RACE/ETHNICITY AND SITE GROUPS FOR COHORT I USING 1986 PRE-TEST SCORES IN SAN DIEGO

Group	Reading Scores in NCEs[a]	Math Scores in NCEs[a]
Overall	-0.49	0.49
Sex		
Males	0.37	-0.09
Females	-1.38	1.13
Race/Ethnicity		
Asians	0.24	1.36
Blacks	-0.01	0.08
Hispanics	-4.51**	-0.79
Whites & Others	4.13	4.88
Site	###	
Boston	1.12	0.70
Fresno	0.42	2.41
San Diego	-6.75***	-2.62
Seattle	3.35*	1.36
Control Group Mean	45.79	52.53
Sample Size	764	759

***Significant at the .01 level.
**Significant at the .05 level.
*Significant at the .10 level.

###Indicates that the impact differs with respect to this characteristic at a .01 level of significance.
##Indicates that the impact differs with respect to this characteristic at a .05 level of significance.
#Indicates that the impact differs with respect to this characteristic at a .10 level of significance.

[a] The test score regressions exclude Portland's test score data because they were not comparable to data from the other sites.

Table C.21

CHANGES IN MEAN READING AND MATH SCORES OVER THE 15-MONTH PROGRAM FOR RETURNING COHORT I TREATMENT YOUTH BY SITE

Site (Sample Size)	Reading Scores	Math Scores
Boston (65)		
1985 Pre-test	732.02	688.82
First Summer Change	-13.64**	4.36
School-year Change	7.78	7.11
Second Summer Change	5.26	3.00
15-Month Change	-0.58	14.47
Fresno (97)		
1985 Pre-test	706.30	692.56
First Summer Change	0.05	-3.69
School-year Change	12.33**	24.15***
Second Summer Change	27.32***	-4.79
15-Month Change	-14.94***	15.67*
Portland (93)		
1985 Pre-test	716.92	682.57
First Summer Change	-4.55	8.91
School-year Change	23.85***	14.09*
Second Summer Change	-26.60***	-5.27
15-Month Change	-7.30	17.73**
San Diego (75)		
1985 Pre-test	729.93	733.79
First Summer Change	-12.26*	-12.94
School-year Change	20.85***	22.30**
Second Summer Change	-8.51	6.44
15-Month Change	0.08	15.80*
Seattle (80)		
1985 Pre-test	699.46	670.09
First Summer Change	-0.16	9.55
School-year Change	24.66***	34.97***
Second Year Change	-19.66***	-17.43**
15-Month Change	4.84	27.09***

***Indicates that change statistically differs from zero at the 0.01 level.

**Indicates that change statistically differs from zero at the 0.05 level.

*Indicates that change statistically differs from zero at the 0.10 level.

APPENDIX D

TOTAL COST OF STEP IN 1986

Table D.1

TOTAL COST OF STEP IN 1986 BY SITE AND FOR THE DEMONSTRATION AS A WHOLE

Program Component	Boston	Fresno	Portland	San Diego	Seattle	Demonstration Average (Including P/PV Costs)
Remediation						
Summer Instruction	$126,151	$114,097	$130,897	$209,434	$159,437	$759,643[a]
Curriculum Development	22,169	19,454	7,402	7,698	12,456	92,428[b]
Total	148,320	133,551	138,299	217,132	171,893	852,071[c]
Life Skills and Opportunities	11,074	9,609	13,499	13,877	16,382	82,941[d]
Work Experience	0	729	3,957	0	1,895	6,581
Other Activities and Administration						
Eligibility Determination	514	109	0	0	1,773	2,396
Re-Enrollment	11,413	2,484	16,786	1,095[e]	9,240	41,018
Program Administration and Support Services	2,426	6,435	24,802	13,467	15,530	60,235
Total	14,353	9,028	41,588	14,562	26,543	103,649
School-Year Support	19,512	45,081	28,579	0	26,500	119,671
Total	$193,259	$197,998	$225,922	$245,571	$243,213	$1,164,913

NOTE: Boston enrolled 153 treatment youth in Cohort I, of which 88 returned in 1986, joining the 158 Cohort II treatment youth. Fresno enrolled 150 treatment youth in Cohort I, of which 124 returned in 1986, joining the 160 Cohort II treatment youth. Portland enrolled 146 treatment youth in Cohort I, of which 122 returned and joined the 155 Cohort II treatment youth. San Diego enrolled 145 treatment youth in Cohort I, of which 109 returned to join the 147 Cohort II youth. Seattle enrolled 158 treatment youth in Cohort I, of which 121 returned, joining the 145 Cohort II treatment youth.

[a] P/PV costs of $23,250 are included.
[b] P/PV costs of $17,200 are included.
[c] P/PV costs of $40,450 are included.
[d] P/PV costs of $18,501 are included.
[e] San Diego re-enrolled both treatments and controls for Cohort I. This $1,095 is the estimated differential cost of re-enrolling the treatments (compared with controls).

APPENDIX E

TECHNICAL DATA TABLES

Table E.1

VARIABLE DESCRIPTIONS

Variable	Description	Type
ABSMUCH	Participant was absent 11 or more days during the previous school year.	Dummy (0,1)
ABSMUCH1	Participant was absent 11 or more days during the previous school year.	Dummy (0,1)
AGE	1 if participant was 14 or under. 0 otherwise.	Dummy (0,1)
ASIAN	Self-explanatory	Dummy (0,1)
ASIAN_T	ASIAN * TREAT interaction term.	
ASIAN_DB	ASIAN interacted with dummy (DB) where DB=1 if student is in the Boston, San Diego or Seattle site, 0 otherwise.	Index
ASIAN_DF	ASIAN interaced with dummy (DF) where DF=1 if student is in the Fresno or Portland site, 0 otherwise.	Index
AVAILBL1	Participants' knowledge of the availability of birth control (index based on number of correct answers to eight availability questions on the pre-test).	Index
AVAILBL2	Participants' knowledge of the availability of birth control (index based on number of correct answers to eight availability questions on the post-test).	Index
BABSM	DB (see Asian_DB) * ABSMUCH interaction term.	Dummy (0,1)
BBRTH	DB (see Asian_DB) * BRTHORDR interaction term	Categorical
BEHS1	DB (see Asian_DB) * EPOSTHS1 interaction term.	Categorical
BENJM	DB (see Asian_DB) * ENJMATH1 interaction term.	Categorical

Table E.1 - (cont'd)

VARIABLE DESCRIPTIONS

Variable	Description	Type
BENJR	DB (see Asian_DB) * ENJREAD1 interaction term.	Categorical
BFAMS	DB (see Asian_DB) * FAMSIZE interaction term.	Categorical
BGR56	DB (see Asian_DB) * GRADE56 interaction term.	Categorical
BLACK	Self-explanatory.	Dummy (0,1)
BLACK_DB	DB (see Asian_DB) * BLACK interaction term.	Dummy (0,1)
BLACK_DF	DF (see Asian_DF) * BLACK interaction term.	Dummy (0,1)
BLACK_T	BLACK * TREAT interaction term.	Dummy (0,1)
BLANP	DB (see Asian_DB) * LANGPROB interaction term.	Dummy (0,1)
BMALE	DB (see Asian_DB) * MALE interaction term.	Dummy (0,1)
BOSTON	Self-explanatory	Dummy (0,1)
BOS_T	BOSTON * TREAT interaction term.	Dummy (0,1)
BOST_TRT	BOSTON * TREAT interaction term.	Dummy (0,1)
BREPE	DB (see Asian_DB) * REPEAT interaction term.	Dummy (0,1)
BRTHORDR	Participants' birth order (e.g., would be three if participant was the third child born in the family).	Categorical
BSTDM	DB (see Asian_DB) * STDMATH1 interaction term.	Continuous
BSTDR	DB (see Asian_DB) * STDREAD1 interaction term.	Continuous

Table E.1 - (cont'd)

VARIABLE DESCRIPTIONS

Variable	Description	Type
CNCONT11	Participant used birth control at last sexual contact, given that the participant had sex in the past two months. Reported at pre-test.	Dummy (0,1)
CNCONTNR	Participant did not respond to contraceptive use question, given that participant had sex in the past two months. Non-response coded as "1", yes and no as "0". Reported at pre-test	Dummy (0,1)
CNCONTR2	Participant used birth control at last sexual contact, given that the participant had sex in the past two months. Reported at post-test.	Dummy (0,1)
ENJMATH1	Measure of math enjoyment from 1 (very much) to 5 (not at all). Reported at pre-test.	Categorical
ENJREAD1	Measure of reading enjoyment from 1 (very much) to 5 (not at all). Reported at pre-test.	Categorical
EPOSTHS1	Participant expects to obtain at least some post-high school education. Reported at pre-test.	Dummy (0,1)
EVERSX11	Has participant ever had sex? Reported at pre-test.	Dummy (0,1)
EVSXNR1	Has participant ever had sex? Non-response coded as "1", yes and no as "0". Reported at pre-test.	Dummy (0,1)
EVERSEX2	Has participant ever had sex? Reported at post-test.	Dummy (0,1)
FABSM	DF (see Asian_DF) * ABSMUCH interaction term.	Dummy (0,1)
FAMSIZE	Number of members in household including participant.	Categorical

Table E.1 - (cont'd)

VARIABLE DESCRIPTIONS

Variable	Description	Type
FAMSIZE1	Number of members in household including participant.	Categorical
FBRTH	DF (see Asian_DF) * BRTHORDR interaction term.	Categorical
FEHS1	DF (see Asian_DF) * EPOSTHS1 interaction term.	Categorical
FEMHDNR	Is participant's household run by a female? Non-response coded as "1", yes and no coded as "0".	Dummy (0,1)
FEMHD11	Is participant's household run by a female?	Dummy (0,1)
FENJM	DF (see Asian_DF) * ENJMATH1 interaction term.	Categorical
FENJR	DF (see Asian_DF) * ENJREAD1 interaction term.	Categorical
FFAMS	DF (see Asian_DF) * FAMSIZE interaction term.	Categorical
FGR56	DF (see Asian_DF) * GRADE56 interaction term.	Categorical
FLANP	DF (see Asian_DF) * LANGPROB interaction term.	Dummy (0,1)
FMALE	DF (see Asian_DF) * MALE interaction term.	Dummy (0,1)
FREPE	DF (see Asian-DF) * REPEAT interaction term.	Dummy (0,1)
FRESNO	Self-explanatory.	Dummy (0,1)
FRES_T	FRESNO * TREAT interaction term.	Dummy (0,1)
FRES_TRT	FRESNO * TREAT interaction term.	Dummy (0,1)
FSTDM	DF (see Asian_DF) * STDMATH1 interaction term.	Continuous

Table E.1 - (cont'd)

VARIABLE DESCRIPTIONS

Variable	Description	Type
FSTDR	DF (see Asian_DF) * STDREAD1 interaction term.	Continuous
GRADE1	Grade in 1985-86 school year.	Categorical
GRADE56	Grade in 1985-86 school year.	Categorical
HISPANIC	Self-explanatory.	Dummy (0,1)
HISP_DB	DB (see Asian_DB) * HISPANIC interaction term.	Dummy (0,1)
HISP_DF	DF (see Asian_DF) * HISPANIC interaction term.	Dummy (0,1)
HISP_T	HISPANIC * TREAT interaction term.	Dummy (0,1)
JSPNSBL1	One if participant thinks both father and mother are responsible for birth control, "0" otherwise. Reported at pre-test.	Dummy (0,1)
KNOWCON1	Two if participant knows consequences of teenage pregnancy to mother and the cost of raising a child to age 18. One if participant knows either. Zero if participant knows neither. Reported at pre-test.	Index
KNOWCON2	As above at 1985 pre-test.	Index
KNOWLDG1	Participant's knowledge of birth control methods (index based on number of correct answers to control usage questions on the pre-test).	Index
KNOWLDG2	As above (reported on the post-test).	Index
LANGPROB	Participant speaks a language in addition to English and indicated problems speaking or understanding English.	Dummy (0,1)

Table E.1 - (cont'd)

VARIABLE DESCRIPTIONS

Variable	Description	Type
LANGPROB1	Participant speaks a language in addition to English and indicated problems speaking or understanding English.	Dummy (0,1)
MALE	Self-explanatory.	Dummy (0,1)
MALE_T	MALE * TREAT interaction term.	Dummy (0,1)
MNCE	Normal curve equivalent to participant's math test score.	Continuous
MNCE3	As above except in San Diego where STDMATH3 was used as the base for constructing the NCE.	Continuous
PCTABS	Percent of time participant was absent during the school year.	Continuous
PCTCRED	Percent of possible credits earned by the participant.	Continuous
PORTLAND	Self-explanatory.	Dummy (0,1)
PORT_T	PORTLAND * TREAT interaction term.	Dummy (0,1)
PORT_TRT	PORTLAND * TREAT interaction term.	Dummy (0,1)
PROMOTED	Dummy for whether participant passed or failed.	Dummy (0,1)
RCNTSEXB	Has participant had sex in the last two months. Reported at post-test.	Dummy (0,1)
REPEAT	Repeated a grade.	Dummy (0,1)
REPEAT1	Repeated a grade.	Dummy (0,1)
RNCE	Normal curve equivalent to participant's reading test score.	Continuous
RNCE3	As above except in San Diego where STDREAD3 was used as a base for constructing the NCE.	Continuous

Table E.1 - (cont'd)

VARIABLE DESCRIPTIONS

Variable	Description	Type
SANDIEGO	Self-explanatory.	Dummy (0,1)
SAND-TRT	SANDIEGO * TREAT interaction term.	Dummy (0,1)
SEXEDNR	Dummy for whether participant previously took a sex education course. Non-response coded as "1", yes and no as "0".	Dummy (0,1)
SEXED11	Dummy for whether participant previously took a sex education course.	Dummy (0,1)
STDMATH1	Math score (scaled) from MAT Survey Test intermediate level. Pre-test (1st summer).	Continuous
STDMATH2	Math score (scaled) from MAT Survey Test intermediate level. Post-test (1st summer).	Continuous
STDMATH3	Math score (scaled) from MAT Survey Test intermediate level. Pre-test (2nd summer).	Continuous
STDMATH4	Math score (scaled) from MAT Survey Test intermediate level. Post-test (2nd summer).	Continuous
STDREAD1	Reading score (scaled) for MAT Survey Test intermediate level. Pre-test (1st summer).	Continuous
STDREAD2	Reading score (scaled) for MAT Survey Test intermediate level. Post-test (1st summer).	Continuous
STDREAD3	Reading score (scaled) for MAT Survey Test intermediate level. Pre-test (2nd summer).	Continuous
STDREAD4	Reading score (scaled) for MAT Survey Test intermediate level. Post-test (2nd summer).	Continuous
TREAT	Treatment or control group.	Dummy (0,1)

Table E.1 - (cont'd)

VARIABLE DESCRIPTIONS

Variable	Description	Type
TREMHRS	Total remediation and life skills hours (1st summer).	Continous
TREM86	Total remediation and life skills hours (2nd summer).	Continous

Table E.2

REGRESSION ANALYSIS OF POST-TEST READING SCORES FOR COHORT II

```
DEP VARIABLE: STDREAD2
ANALYSIS OF VARIANCE

                        SUM OF          MEAN
SOURCE      DF         SQUARES        SQUARE        F VALUE        PROB>F

MODEL       18      2822507.08     156805.95         70.504        0.0001
ERROR     1249      2777879.42    2224.08281
C TOTAL   1267      5600386.51

    ROOT MSE          47.16018      R-SQUARE         0.5040
    DEP MEAN          709.4803      ADJ R-SQ         0.4968
    C.V.              6.647145

PARAMETER ESTIMATES

                       PARAMETER        STANDARD        T FOR HO:
VARIABLE    DF          ESTIMATE           ERROR      PARAMETER=0        PROB > |T|

INTERCEP     1         148.28519     24.41787222            6.073            0.0001
TREAT        1       11.79015147      2.65931574            4.434            0.0001
MALE         1       -9.13124764      2.74478125           -3.327            0.0009
BLACK        1       -9.06361367      4.04652417           -2.240            0.0253
ASIAN        1       13.43354827      6.11499401            2.197            0.0282
HISPANIC     1        4.89723045      5.41884022            0.904            0.3663
FAMSIZE1     1       -0.31306858      0.69154822           -0.453            0.6508
BRTHORDR     1        1.14243221      0.84118161            1.358            0.1747
GRADE1       1        3.73337503      2.00644906            1.861            0.0630
REPEAT1      1       -8.55278862      3.19953665           -2.673            0.0076
ABSMUCH1     1       -4.84591731      3.31324496           -1.463            0.1438
EPOSTHS1     1        5.79251216      3.09690374            1.870            0.0617
LANGPRB1     1       -4.57907216      4.54628365           -1.007            0.3140
ENJREAD1     1       -2.74859043      1.32344233           -2.077            0.0380
STDREAD1     1        0.73015556      0.02404386           30.368            0.0001
BOSTON       1       21.25332817      4.44277785            4.784            0.0001
FRESNO       1        2.63207646      4.54969897            0.579            0.5630
PORTLAND     1        3.94822683      4.34210291            0.909            0.3634
SANDIEGO     1       16.42494679      4.32870214            3.794            0.0002
```

Table E.3

REGRESSION ANALYSIS OF POST-TEST MATH SCORES FOR COHORT II

```
DEP VARIABLE: STDMATH2
ANALYSIS OF VARIANCE
                        SUM OF          MEAN
SOURCE      DF         SQUARES        SQUARE        F VALUE       PROB>F

MODEL       19      5592821.23     294359.01         82.420       0.0001
ERROR     1248      4457160.72    3571.44289
C TOTAL   1267     10049981.95

    ROOT MSE          59.76155      R-SQUARE         0.5565
    DEP MEAN          697.9937      ADJ R-SQ         0.5497
    C.V.              8.561904

PARAMETER ESTIMATES

                       PARAMETER           STANDARD       T FOR HO:
VARIABLE    DF          ESTIMATE              ERROR     PARAMETER=0         PROB > |T|

INTERCEP     1       69.11830164        31.05361872           2.226             0.0262
TREAT        1       23.13455544         3.37250972           6.860             0.0001
MALE         1      -11.92932569         3.45676251          -3.451             0.0006
BLACK        1      -11.01361868         5.15257976          -2.137             0.0328
ASIAN        1       23.49557046         7.84947540           2.993             0.0028
HISPANIC     1       -2.28903894         6.86113264          -0.334             0.7387
FAMSIZE1     1       -1.33253352         0.87597338          -1.521             0.1285
BRTHORDR     1        0.27800435         1.06779380           0.260             0.7946
GRADE1       1        4.68010196         2.55047984           1.835             0.0667
REPEAT1      1      -10.24500104         4.06154307          -2.522             0.0118
ABSMUCH1     1        1.72114092         4.21153269           0.409             0.6828
EPOSTHS1     1        3.79448338         3.94971055           0.961             0.3369
LANGPRB1     1       -1.83224227         5.75412841          -0.318             0.7502
ENJMATH1     1       -8.59132231         1.56463220          -5.491             0.0001
STDMATH1     1        0.62475653         0.03092233          20.204             0.0001
STDREAD1     1        0.21762814         0.03568883           6.098             0.0001
BOSTON       1       28.27718505         5.63902162           5.015             0.0001
FRESNO       1       12.70029392         5.78263511           2.196             0.0283
PORTLAND     1       10.48045351         5.51027619           1.902             0.0574
SANDIEGO     1       32.24022150         5.54688371           5.812             0.0001
```

Table E.4

REGRESSION ANALYSIS OF POST-TEST READING SCORES FOR COHORT II WITH SITE-TREATMENT INTERACTION TERMS

```
DEP VARIABLE: STDREAD2
ANALYSIS OF VARIANCE

                      SUM OF         MEAN
SOURCE      DF       SQUARES       SQUARE        F VALUE        PROB>F

MODEL       22    2829822.63    128628.30         57.801        0.0001
ERROR     1245    2770563.88   2225.35251
C TOTAL   1267    5600386.51

    ROOT MSE        47.17364     R-SQUARE         0.5053
    DEP MEAN        709.4803     ADJ R-SQ         0.4965
    C.V.            6.649042

PARAMETER ESTIMATES

                     PARAMETER        STANDARD      T FOR HO:
VARIABLE   DF         ESTIMATE           ERROR    PARAMETER=0      PROB > |T|

INTERCEP    1        142.53989     24.67821507          5.776          0.0001
TREAT       1      20.26578065      6.04902990          3.350          0.0008
MALE        1      -9.02814175      2.74750685         -3.286          0.0010
BLACK       1      -8.54436477      4.05935065         -2.105          0.0355
ASIAN       1      13.68907056      6.12120728          2.236          0.0255
HISPANIC    1       4.98931164      5.42518802          0.920          0.3579
FAMSIZE1    1      -0.30096952      0.69200393         -0.435          0.6637
BRTHORDR    1       1.08545461      0.84250145          1.288          0.1979
GRADE1      1       3.88355690      2.00912368          1.933          0.0535
REPEAT1     1      -8.27087492      3.20494793         -2.581          0.0100
ABSMUCH1    1      -5.14119319      3.31930481         -1.549          0.1217
EPOSTHS1    1       5.71409834      3.10001422          1.843          0.0655
LANGPRB1    1      -4.31518777      4.55423281         -0.948          0.3436
ENJREAD1    1      -2.66435939      1.32640696         -2.009          0.0448
STDREAD1    1       0.72961653      0.02408502         30.293          0.0001
BOSTON      1      25.58333086      6.11875320          4.181          0.0001
FRESNO      1      10.18706183      6.24307460          1.632          0.1030
PORTLAND    1       8.04028258      6.16720826          1.304          0.1926
SANDIEGO    1      21.80872199      6.07955941          3.587          0.0003
BOST_TRT    1      -8.49039063      8.42285335         -1.008          0.3136
FRES_TRT    1     -14.94408576      8.48506492         -1.761          0.0784
PORT_TRT    1      -7.90794465      8.53582619         -0.926          0.3544
SAND_TRT    1     -10.67570713      8.45431287         -1.263          0.2069
```

Table E.5

REGRESSION ANALYSIS OF POST-TEST MATH SCORES FOR COHORT II WITH SITE-TREATMENT INTERACTION TERMS

```
DEP VARIABLE: STDMATH2
ANALYSIS OF VARIANCE

                      SUM OF         MEAN
SOURCE     DF        SQUARES       SQUARE        F VALUE       PROB>F

MODEL      23     5621129.78    244396.95         68.648       0.0001
ERROR    1244     4428852.16   3560.17055
C TOTAL  1267    10049981.95

    ROOT MSE        59.66716     R-SQUARE         0.5593
    DEP MEAN        697.9937     ADJ R-SQ         0.5512
    C.V.            8.548382

PARAMETER ESTIMATES

                        PARAMETER         STANDARD      T FOR HO:
VARIABLE  DF             ESTIMATE            ERROR    PARAMETER=0        PROB > |T|

INTERCEP   1          57.87862830      31.31177900          1.848            0.0648
TREAT      1          41.79673022       7.65063239          5.463            0.0001
MALE       1         -11.59956464       3.45378733         -3.359            0.0008
BLACK      1          -9.94129728       5.15922830         -1.927            0.0542
ASIAN      1          24.32159424       7.84290888          3.101            0.0020
HISPANIC   1          -1.81219480       6.85602350         -0.264            0.7916
FAMSIZE1   1          -1.30915105       0.87493094         -1.496            0.1348
BRTHORDR   1           0.21127571       1.06727684          0.198            0.8431
GRADE1     1           4.92431484       2.54930579          1.932            0.0536
REPEAT1    1          -9.71330026       4.06071659         -2.392            0.0169
ABSMUCH1   1           1.22264608       4.21136740          0.290            0.7716
EPOSTHS1   1           3.66400607       3.94602472          0.929            0.3533
LANGPRB1   1          -1.63850395       5.75273456         -0.285            0.7758
ENJMATH1   1          -8.48261854       1.56404150         -5.424            0.0001
STDMATH1   1           0.62535835       0.03091876         20.226            0.0001
STDREAD1   1           0.21503524       0.03567093          6.028            0.0001
BOSTON     1          39.71145529       7.74809138          5.125            0.0001
FRESNO     1          23.34208335       7.91412453          2.949            0.0032
PORTLAND   1          21.44549081       7.81221821          2.745            0.0061
SANDIEGO   1          46.34243062       7.73094759          5.994            0.0001
BOST_TRT   1         -22.58430559      10.65508542         -2.120            0.0342
FRES_TRT   1         -20.83082635      10.72783009         -1.942            0.0524
PORT_TRT   1         -21.25491450      10.80484524         -1.967            0.0494
SAND_TRT   1         -27.93277804      10.69637521         -2.611            0.0091
```

Table E.6

TWO-STAGE LEAST SQUARES ANALYSIS OF POST-TEST READING SCORES FOR COHORT II WITH TREATMENT HOURS

```
MODEL: EQU  1      SECOND STAGE
DEP VARIABLE: STDREAD2
ANALYSIS OF VARIANCE

                      SUM OF          MEAN
SOURCE      DF       SQUARES        SQUARE        F VALUE       PROB>F

MODEL       18    2822507.08     156805.95         70.758       0.0001
ERROR     1249    2767881.19    2216.07781
C TOTAL   1267    5600386.51

     ROOT MSE       47.07524      R-SQUARE         0.5049
     DEP MEAN       709.4803      ADJ R-SQ         0.4978
     C.V.           6.635172

PARAMETER ESTIMATES

                         PARAMETER          STANDARD        T FOR HO:
VARIABLE    DF            ESTIMATE             ERROR      PARAMETER=0          PROB > |T|

INTERCEP     1           150.77510       24.35374760            6.191              0.0001
TREMHRS      1          0.11931442        0.02686337            4.442              0.0001
MALE         1         -9.02446098        2.74087014           -3.293              0.0010
BLACK        1         -9.41892836        4.04163529           -2.330              0.0199
HISPANIC     1          4.54879746        5.41154515            0.841              0.4007
ASIAN        1         12.76862770        6.10953553            2.090              0.0368
FAMSIZE1     1         -0.34850984        0.69065387           -0.505              0.6139
BRTHORDR     1          1.18030896        0.83971753            1.406              0.1601
GRADE1       1          3.79184338        2.00238946            1.894              0.0585
REPEAT1      1         -8.56425330        3.19380374           -2.682              0.0074
ABSMUCH1     1         -4.15943193        3.30738757           -1.258              0.2088
EPOSTHS1     1          5.57785393        3.09023620            1.805              0.0713
LANGPRB1     1         -4.95941943        4.53755248           -1.093              0.2746
ENJREAD1     1         -2.82192444        1.32090929           -2.136              0.0328
STDREAD1     1          0.72774370        0.02399118           30.334              0.0001
BOSTON       1         19.68674679        4.44972690            4.424              0.0001
FRESNO       1          1.62372121        4.54423211            0.357              0.7209
PORTLAND     1          4.28239669        4.33404589            0.988              0.3233
SANDIEGO     1         15.83019543        4.32231725            3.662              0.0003
```

Table E.7

TWO-STAGE LEAST SQUARES ANALYSIS OF POST-TEST MATH SCORES FOR COHORT II WITH TREATMENT HOURS

```
MODEL: EQU  1     SECOND STAGE
DEP VARIABLE: STDMATH2
ANALYSIS OF VARIANCE

                      SUM OF          MEAN
SOURCE      DF       SQUARES        SQUARE        F VALUE       PROB>F

MODEL       19    5592821.23     294359.01         82.249       0.0001
ERROR     1248    4466456.04    3578.89106
C TOTAL   1267   10049981.95

     ROOT MSE       59.82383      R-SQUARE         0.5560
     DEP MEAN       697.9937      ADJ R-SQ         0.5492
     C.V.           8.570827

PARAMETER ESTIMATES

                        PARAMETER          STANDARD       T FOR HO:
VARIABLE   DF            ESTIMATE             ERROR     PARAMETER=0       PROB > |T|

INTERCEP    1         74.06179797       31.06059649           2.384           0.0173
TREMHRS     1          0.23394640        0.03413979           6.853           0.0001
MALE        1        -11.73351383        3.46157355          -3.390           0.0007
BLACK       1        -11.71492536        5.16108007          -2.270           0.0234
HISPANIC    1         -2.99361001        6.87147025          -0.436           0.6632
ASIAN       1         22.41737760        7.86428491           2.851           0.0044
FAMSIZE1    1         -1.39595678        0.87730063          -1.591           0.1118
BRTHORDR    1          0.35981840        1.06896303           0.337           0.7365
GRADE1      1          4.82521752        2.55237470           1.890           0.0589
REPEAT1     1        -10.31527480        4.06587609          -2.537           0.0113
ABSMUCH1    1          3.03366182        4.21648303           0.719           0.4720
EPOSTHS1    1          3.43501703        3.95285116           0.869           0.3850
LANGPRB1    1         -2.51980785        5.75968506          -0.437           0.6618
STDREAD1    1          0.21588978        0.03572524           6.043           0.0001
ENJMATH1    1         -8.56691998        1.56620950          -5.470           0.0001
STDMATH1    1          0.62033422        0.03094065          20.049           0.0001
BOSTON      1         25.25763296        5.66411690           4.459           0.0001
FRESNO      1         10.80615817        5.79263683           1.865           0.0623
PORTLAND    1         11.18659324        5.51551308           2.028           0.0428
SANDIEGO    1         31.21418508        5.55493858           5.619           0.0001
```

Table E.8

REGRESSION ANALYSIS OF POST-TEST CONTRACEPTIVE KNOWLEDGE FOR COHORT II

```
DEP VARIABLE: KNOWLDG2
ANALYSIS OF VARIANCE

                      SUM OF          MEAN
SOURCE      DF       SQUARES        SQUARE        F VALUE        PROB>F

MODEL       25    5727.95641     229.11826         49.998        0.0001
ERROR     1276    5847.32240    4.58254106
C TOTAL   1301   11575.27880

    ROOT MSE        2.140687      R-SQUARE         0.4948
    DEP MEAN        6.430876      ADJ R-SQ         0.4849
    C.V.            33.28765

PARAMETER ESTIMATES

                        PARAMETER        STANDARD       T FOR HO:
VARIABLE    DF           ESTIMATE           ERROR     PARAMETER=0       PROB > |T|

INTERCEP     1        -3.26828303      0.86409090          -3.782          0.0002
TREAT        1         2.64667416      0.11978680          22.095           0.0001
MALE         1        -0.04244354      0.13055914          -0.325           0.7452
AGE          1        -0.23981860      0.12175855          -1.970           0.0491
BLACK        1        -0.08542551      0.17960568          -0.476           0.6344
HISPANIC     1        -0.07225086      0.24334454          -0.297           0.7666
ASIAN        1         0.05044397      0.27835319           0.181           0.8562
FAMSIZE1     1       -0.006720325      0.03277588          -0.205           0.8376
FEMHD11      1         0.09232412      0.13703154           0.674           0.5006
FEMHDNR      1        -0.10551488      0.28530912          -0.370           0.7116
BRTHORDR     1        0.003186733      0.03761383           0.085           0.9325
LANGPRB1     1        -0.29699850      0.20567941          -1.444           0.1490
ABSMUCH1     1        0.009746228      0.14886714           0.065           0.9478
SEXED11      1         0.60138684      0.14467532           4.157           0.0001
SEXEDNR      1         0.46755954      0.27994109           1.670           0.0951
EVERSX11     1         0.18820507      0.14852009           1.267           0.2053
EVSXNR1      1        -0.07227155      0.24561227          -0.294           0.7686
STDREAD1     1        0.006847255     0.001122852           6.098           0.0001
AVAILBL1     1         0.19912441      0.03852805           5.168           0.0001
KNOWLDG1     1         0.37333278      0.03155830          11.830           0.0001
JSPNSBL1     1         0.24013929      0.14192876           1.692           0.0909
KNOWCON1     1        -0.07263592      0.10034923          -0.724           0.4693
BOSTON       1         0.18872633      0.19570239           0.964           0.3351
FRESNO       1         0.79952455      0.20469845           3.906           0.0001
PORTLAND     1         0.63435574      0.19410861           3.268           0.0011
SANDIEGO     1         1.40077045      0.19735818           7.098           0.0001
```

Table E.9

REGRESSION ANALYSIS OF POST-TEST KNOWLEDGE OF AVAILABILITY OF CONTRACEPTION FOR COHORT II

```
DEP VARIABLE: AVAILBL2
ANALYSIS OF VARIANCE

                       SUM OF         MEAN
SOURCE      DF        SQUARES       SQUARE        F VALUE       PROB>F

MODEL       25     2532.96367    101.31855         35.402       0.0001
ERROR     1276     3651.80745   2.86191806
C TOTAL   1301     6184.77112

     ROOT MSE         1.69172     R-SQUARE         0.4095
     DEP MEAN        4.302611     ADJ R-SQ         0.3980
     C.V.            39.31846

PARAMETER ESTIMATES

                        PARAMETER          STANDARD       T FOR H0:
VARIABLE   DF            ESTIMATE             ERROR     PARAMETER=0        PROB > |T|

INTERCEP    1         -2.41472850        0.68286499          -3.536            0.0004
TREAT       1          1.74087768        0.09466390          18.390            0.0001
MALE        1         -0.21788814        0.10317696          -2.112            0.0349
AGE         1         -0.09862510        0.09622212          -1.025            0.3056
BLACK       1         -0.14539149        0.14193695          -1.024            0.3059
HISPANIC    1          0.36278068        0.19230785           1.886            0.0595
ASIAN       1          0.28184437        0.21997414           1.281            0.2003
FAMSIZE1    1          0.01035090        0.02590179           0.400            0.6895
FEMHD11     1          0.21907614        0.10829189           2.023            0.0433
FEMHDNR     1          0.21691330        0.22547119           0.962            0.3362
BRTHORDR    1          0.02399280        0.02972508           0.807            0.4197
LANGPRB1    1         -0.29961621        0.16254223          -1.843            0.0655
ABSMUCH1    1         -0.23652800        0.11764521          -2.011            0.0446
SEXED11     1          0.30898338        0.11433255           2.702            0.0070
SEXEDNR     1          0.29848034        0.22122900           1.349            0.1775
EVERSX11    1         -0.04446860        0.11737095          -0.379            0.7048
EVSXNR1     1         -0.09587699        0.19409997          -0.494            0.6214
STDREAD1    1         0.005037220       0.000887356           5.677            0.0001
AVAILBL1    1          0.31747922        0.03044756          10.427            0.0001
KNOWLDG1    1          0.13489686        0.02493957           5.409            0.0001
JSPNSBL1    1          0.10574534        0.11216202           0.943            0.3460
KNOWCON1    1          0.04717167        0.07930297           0.595            0.5521
BOSTON      1          0.11006255        0.15465770           0.712            0.4768
FRESNO      1          0.23930748        0.16176701           1.479            0.1393
PORTLAND    1          0.18529849        0.15339818           1.208            0.2273
SANDIEGO    1          0.76905704        0.15596622           4.931            0.0001
```

Table E.10

REGRESSION ANALYSIS OF POST-TEST KNOWLEDGE OF THE CONSEQUENCES OF ADOLESCENT PREGNANCY FOR COHORT II

```
DEP VARIABLE: KNOWCON2
ANALYSIS OF VARIANCE

                      SUM OF          MEAN
SOURCE      DF       SQUARES        SQUARE        F VALUE       PROB>F

MODEL       23   88.95651704    3.86767465         10.461       0.0001
ERROR     1275     471.39606    0.36972240
C TOTAL   1298     560.35258

    ROOT MSE        0.608048      R-SQUARE         0.1588
    DEP MEAN        1.022325      ADJ R-SQ         0.1436
    C.V.            59.47699

PARAMETER ESTIMATES

                        PARAMETER        STANDARD      T FOR HO:
VARIABLE    DF           ESTIMATE           ERROR    PARAMETER=0       PROB > |T|

INTERCEP     1        -0.38449662      0.24304369         -1.582           0.1139
TREAT        1         0.25212945      0.03402041          7.411           0.0001
MALE         1        -0.02616539      0.03707046         -0.706           0.4804
AGE          1         0.08145017      0.03459876          2.354           0.0187
BLACK        1        -0.04500615      0.05106585         -0.881           0.3783
HISPANIC     1       -0.008549527      0.06922598         -0.124           0.9017
ASIAN        1         0.02650839      0.07843609          0.338           0.7354
FAMSIZE1     1       -0.001877433     0.009295832         -0.202           0.8400
FEMHD11      1        0.000264046      0.03880552          0.007           0.9946
FEMHDNR      1         0.06348096      0.08147843          0.779           0.4361
BRTHORDR     1        0.006829250      0.01065004          0.641           0.5215
LANGPRB1     1         0.06114310      0.05844746          1.046           0.2957
ABSMUCH1     1        0.005145071      0.04231705          0.122           0.9032
EVERSX11     1        -0.07473115      0.04175135         -1.790           0.0737
EVSXNR1      1         0.02974276      0.06975789          0.426           0.6699
STDREAD1     1        0.001286670     0.000311641          4.129           0.0001
SEXED11      1        -0.02067980      0.04069186         -0.508           0.6114
SEXEDNR      1        -0.07642675      0.07950874         -0.961           0.3366
KNOWCON1     1         0.26502236      0.02849065          9.302           0.0001
JSPNSBL1     1         0.01261904      0.03999455          0.316           0.7524
BOSTON       1         0.05700892      0.05553282          1.027           0.3048
FRESNO       1         0.15597740      0.05814841          2.682           0.0074
PORTLAND     1         0.04051753      0.05514996          0.735           0.4627
SANDIEGO     1         0.22477216      0.05609238          4.007           0.0001
```

Table E.11

LOGIT ANALYSIS OF THE PROBABILITY OF BEING SEXUALLY EXPERIENCED AT POST-TEST FOR COHORT II

```
                         DEPENDENT VARIABLE: EVERSEX2

          1261 OBSERVATIONS
           619  EVERSEX2=  0
           642  EVERSEX2=  1
           374 OBSERVATIONS DELETED DUE TO MISSING VALUES

VARIABLE          BETA           STD. ERROR       CHI-SQUARE     P         R

INTERCEPT      3.98893028        1.78524360          4.99    0.0255
TREAT         -0.49722044        0.23246977          4.57    0.0324   -0.038
AGE           -0.42359843        0.23326237          3.30    0.0694   -0.027
BLACK          0.30357357        0.30570799          0.99    0.3207    0.000
HISPANIC       0.76792832        0.36858462          4.34    0.0372    0.037
MALE           1.14924254        0.24578857         21.86    0.0000    0.107
FAMSIZE1      -0.00165876        0.06516753          0.00    0.9797    0.000
FEMHD11        0.37555631        0.25852700          2.11    0.1463    0.008
FEMHDNR       -0.18714888        0.56525025          0.11    0.7406    0.000
BRTHORDR      -0.21173054        0.07630673          7.70    0.0055   -0.057
LANGPRB1      -1.22131283        0.32940367         13.75    0.0002   -0.082
ABSMUCH1       0.16281553        0.28995355          0.32    0.5744    0.000
STDREAD1      -0.00805729        0.00235069         11.75    0.0006   -0.075
EVERSX11       6.69606162        0.53873798        154.48      .       0.295
EVSXNR1        2.76925124        0.31456575         77.50      .       0.208
SEXED11        0.31323582        0.27821078          1.27    0.2602    0.000
SEXEDNR        0.34996531        0.54779081          0.41    0.5229    0.000
AVAILBL1       0.10700892        0.07437989          2.07    0.1502    0.006
KNOWLDG1      -0.05098671        0.05886345          0.75    0.3864    0.000
JSPNSBL1       0.00109390        0.26746980          0.00    0.9967    0.000
BOSTON         0.16658697        0.37155625          0.20    0.6539    0.000
FRESNO        -0.19635728        0.39968357          0.24    0.6232    0.000
PORTLAND       0.07911409        0.37605987          0.04    0.8334    0.000
SANDIEGO       0.20447898        0.39381980          0.27    0.6036    0.000
```

Table E.12

LOGIT ANALYSIS OF THE PROBABILITY OF BEING SEXUALLY ACTIVE DURING THE STEP PROGRAM FOR COHORT II

```
                         DEPENDENT VARIABLE: RCNTSEXB

         1208 OBSERVATIONS
          871  RCNTSEXB=  0
          337  RCNTSEXB=  1
          427 OBSERVATIONS DELETED DUE TO MISSING VALUES

VARIABLE          BETA          STD. ERROR      CHI-SQUARE    P        R

INTERCEPT   -0.80365231        1.19783544         0.45   0.5023
TREAT        0.14578899        0.17284695         0.71   0.3990    0.000
MALE         0.35307114        0.18713146         3.56   0.0592    0.033
AGE         -0.18330480        0.17386694         1.11   0.2918    0.000
BLACK        0.44305960        0.22512853         3.87   0.0491    0.036
HISPANIC     0.42012282        0.29986390         1.96   0.1612    0.000
FAMSIZE1     0.03200766        0.04970210         0.41   0.5196    0.000
FEMHD11      0.20339586        0.19090205         1.14   0.2867    0.000
FEMHDNR      0.47656009        0.41902795         1.29   0.2554    0.000
BRTHORDR    -0.12899152        0.05401371         5.70   0.0169   -0.051
LANGPRB1    -0.64760055        0.25718163         6.34   0.0118   -0.055
ABSMUCH1     0.56252954        0.19657157         8.19   0.0042    0.066
STDREAD1    -0.00326067        0.00153475         4.51   0.0336   -0.042
EVERSX11     3.27854572        0.23749377       190.57    .        0.363
EVSXNR1      1.41149260        0.50583124         7.79   0.0053    0.064
SEXED11     -0.23723724        0.21206781         1.25   0.2633    0.000
SEXEDNR     -0.65275605        0.37531704         3.02   0.0820   -0.027
AVAILBL1     0.12870650        0.05666141         5.16   0.0231    0.047
KNOWLDG1    -0.04167890        0.04585740         0.83   0.3634    0.000
JSPNSBL1    -0.28463448        0.20423025         1.94   0.1634    0.000
BOSTON      -0.13614147        0.27043630         0.25   0.6147    0.000
FRESNO      -0.16692458        0.29918215         0.31   0.5769    0.000
PORTLAND     0.20144308        0.27097137         0.55   0.4572    0.000
SANDIEGO     0.08089870        0.28212043         0.08   0.7743    0.000
```

Table E.13

LOGIT ANALYSIS OF THE PROBABILITY OF USING CONTRACEPTION DURING LAST INTERCOURSE FOR SEXUALLY ACTIVE COHORT II YOUTH

DEPENDENT VARIABLE: CNCONTR2

331 OBSERVATIONS
166 CNCONTR2= 0
165 CNCONTR2= 1
36 OBSERVATIONS DELETED DUE TO MISSING VALUES

VARIABLE	BETA	STD. ERROR	CHI-SQUARE	P	R
INTERCEPT	-1.52423247	1.71910833	0.79	0.3753	
TREAT	0.83322650	0.26567388	9.84	0.0017	0.131
AGE	-0.09732328	0.26083058	0.14	0.7091	0.000
BLACK	0.23482060	0.35675286	0.43	0.5104	0.000
HISPANIC	0.47724149	0.50318859	0.90	0.3429	0.000
MALE	0.93980001	0.29397793	10.22	0.0014	0.134
FAMSIZE1	-0.04974030	0.07540513	0.44	0.5095	0.000
FEMHD11	0.26082817	0.29070031	0.81	0.3696	0.000
FEMHDNR	-0.31651614	0.59063723	0.29	0.5920	0.000
BRTHORDR	-0.00477694	0.08283731	0.00	0.9540	0.000
LANGPRB1	0.20761048	0.43202811	0.23	0.6308	0.000
ABSMUCH1	-0.54376393	0.29688056	3.35	0.0670	-0.054
STDREAD1	0.00134734	0.00231654	0.34	0.5608	0.000
CNCONT11	1.17860100	0.28552120	17.04	0.0000	0.181
CNCONTNR	0.29694051	0.45161370	0.43	0.5109	0.000
SEXED11	-0.04392559	0.30860309	0.02	0.8868	0.000
SEXEDNR	-0.20276700	0.56990162	0.13	0.7220	0.000
AVAILBL1	0.08239101	0.08183838	1.01	0.3141	0.000
KNOWLDG1	-0.03764098	0.06832734	0.30	0.5817	0.000
JSPNSBL1	0.02464520	0.29210840	0.01	0.9328	0.000
BOSTON	-1.24986204	0.42139110	8.80	0.0030	-0.122
FRESNO	-2.18388668	0.53453803	16.69	0.0000	-0.179
PORTLAND	-0.57634612	0.41497195	1.93	0.1649	0.000
SANDIEGO	-0.77566770	0.42333102	3.36	0.0669	-0.054

Table E.14

TWO-STAGE LEAST SQUARES ANALYSIS OF SECOND SUMMER POST-TEST READING SCORES FOR COHORT I TREATMENTS

```
MODEL: EQU  1      SECOND STAGE
DEP VARIABLE: STDREAD4
ANALYSIS OF VARIANCE

                    SUM OF          MEAN
SOURCE     DF      SQUARES        SQUARE         F VALUE        PROB>F

MODEL      10    942248.74   94224.87361          30.055        0.0001
ERROR     386   1210151.53    3135.10758
C TOTAL   396   2118939.88

    ROOT MSE      55.99203      R-SQUARE          0.4378
    DEP MEAN      712.7985      ADJ R-SQ          0.4232
    C.V.           7.85524

PARAMETER ESTIMATES

                       PARAMETER         STANDARD       T FOR HO:
VARIABLE  DF            ESTIMATE            ERROR     PARAMETER=0        PROB > |T|

INTERCEP   1           266.93296      76.12239695           3.507            0.0005
TREM86     1          0.35380263       0.30015575           1.179            0.2392
MALE       1        -18.17160833       6.09977811          -2.979            0.0031
BLACK      1        -12.85120135      12.77315775          -1.006            0.3150
GRADE1     1        -12.64307072       5.85577917          -2.159            0.0315
REPEAT1    1          1.67422921       8.21278778           0.204            0.8386
ABSMUCH1   1          4.07460147      26.03456398           0.157            0.8757
EPOSTHS1   1          5.87213604      17.40939871           0.337            0.7361
LANGPRB1   1          4.35778312      35.36995695           0.123            0.9020
ENJREAD1   1          3.05266095       5.43942715           0.561            0.5750
STDREAD3   1          0.71381062       0.05424252          13.160            0.0001
```

Table E.15

TWO-STAGE LEAST SQUARES ANALYSIS OF SECOND SUMMER POST-TEST MATH SCORES FOR COHORT I TREATMENTS

```
MODEL: EQU  1     SECOND STAGE
DEP VARIABLE: STDMATH4
ANALYSIS OF VARIANCE

                       SUM OF          MEAN
SOURCE      DF        SQUARES        SQUARE        F VALUE         PROB>F

MODEL       11     2203247.35     200295.21         29.873         0.0001
ERROR      386     2588072.41    6704.85081
C TOTAL    397     4434658.25

     ROOT MSE        81.88315      R-SQUARE         0.4598
     DEP MEAN        709.8518      ADJ R-SQ         0.4444
     C.V.            11.53525

PARAMETER ESTIMATES

                         PARAMETER           STANDARD       T FOR HO:
VARIABLE   DF             ESTIMATE              ERROR     PARAMETER=0          PROB > |T|

INTERCEP    1         -58.28358847       106.60761           -0.547              0.5849
TREM86      1           1.05630242       0.48088255           2.197              0.0286
MALE        1         -22.48660696       8.77144792          -2.564              0.0107
BLACK       1          -8.69828275      18.53466142          -0.469              0.6391
GRADE1      1           7.65539503       8.79275917           0.871              0.3845
REPEAT1     1          28.50590324      11.60745607           2.456              0.0145
ABSMUCH1    1          52.12681804      41.12119769           1.268              0.2057
EPOSTHS1    1          52.87081784      26.61499579           1.987              0.0477
LANGPRB1    1          40.12142273      58.43666438           0.687              0.4928
ENJMATH1    1         -17.80397448       6.61087636          -2.693              0.0074
STDREAD3    1           0.34091637       0.08710825           3.914              0.0001
STDMATH3    1           0.48729693       0.10191209           4.782              0.0001
```

Table E.16

REGRESSION ANALYSIS OF READING SCORES (IN NCES) WITH SITE-TREATMENT INTERACTION TERMS (PORTLAND EXCLUDED)

```
DEP VARIABLE: RNCE
ANALYSIS OF VARIANCE

                     SUM OF          MEAN
SOURCE     DF       SQUARES        SQUARE        F VALUE        PROB)F

MODEL      20   89302.40662    4465.12033         39.744        0.0001
ERROR     654   73474.62725  · 112.34652
C TOTAL   674     162777.03

     ROOT MSE      10.59936      R-SQUARE         0.5486
     DEP MEAN      40.62844      ADJ R-SQ         0.5348
     C.V.          26.08853

PARAMETER ESTIMATES

                     PARAMETER          STANDARD      T FOR H0:
VARIABLE    DF        ESTIMATE             ERROR    PARAMETER=0       PROB > |T|

INTERCEP     1      -100.93018        7.85145671        -12.855           0.0001
TREAT        1      3.29107045        1.62340467          2.027           0.0430
BOS_T        1     -1.99447936        2.22705554         -0.896           0.3708
FRESNO_T     1     -3.25061849        2.21685556         -1.466           0.1430
DIEGO_T      1     -1.97503034        2.68738903         -0.735           0.4626
STDREAD1     1      0.17468891       0.007991671         21.859           0.0001
ENJREAD1     1     -0.84007756        0.40470584         -2.076           0.0383
REPEAT       1     -2.70273160        1.03808315         -2.604           0.0094
GRADE56      1      2.29996033        0.65923871          3.489           0.0005
ABSMUCH      1     -0.14706213        1.00280753         -0.147           0.8835
EPOSTHS1     1      1.52629397        0.98681778          1.547           0.1224
FAMSIZE      1     -0.39411611        0.21121473         -1.866           0.0625
BRTHORDR     1     0.009156558        0.18366482          0.050           0.9603
LANGPROB     1     -2.70690217        1.77957504         -1.521           0.1287
MALE         1      2.69167871        0.84818681          3.173           0.0016
BLACK        1      0.19721603        1.56392922          0.126           0.8997
HISPANIC     1     -2.34795867        1.81550356         -1.293           0.1964
ASIAN        1     -0.38141044        2.14658393         -0.178           0.8590
BOSTON       1     -5.09616849        1.65272449         -3.083           0.0021
FRESNO       1     -5.79746424        1.69109092         -3.428           0.0006
SANDIEGO     1     -1.78120016        1.87300704         -0.951           0.3420
```

Table E.17

REGRESSION ANALYSIS OF READING SCORES WITH RACIAL/ETHNIC- AND SEX-TREATMENT INTERACTION TERMS (PORTLAND EXCLUDED)

```
DEP VARIABLE: RNCE
ANALYSIS OF VARIANCE

                      SUM OF           MEAN
SOURCE      DF       SQUARES         SQUARE         F VALUE        PROB>F

MODEL       21   89861.02506     4279.09643          38.321        0.0001
ERROR      653   72916.00881      111.66311
C TOTAL    674     162777.03

     ROOT MSE       10.56708     R-SQUARE          0.5520
     DEP MEAN       40.62844     ADJ R-SQ          0.5376
     C.V.           26.00906

PARAMETER ESTIMATES

                     PARAMETER          STANDARD       T FOR H0:
VARIABLE    DF        ESTIMATE             ERROR     PARAMETER=0         PROB > |T|

INTERCEP     1      -101.77874        7.86127227         -12.947             0.0001
TREAT        1      5.18104017        3.01378935           1.719             0.0861
BLACK_T      1     -2.85696485        3.07684233          -0.929             0.3535
HISP_T       1     -6.66098270        3.45264375          -1.929             0.0541
ASIAN_T      1     -0.37509094        3.59301714          -0.104             0.9169
MALE_T       1     -1.52116845        1.65584993          -0.919             0.3586
STDREAD1     1      0.17502947       0.007953530          22.007             0.0001
ENJREAD1     1     -0.85190280        0.40227544          -2.118             0.0346
REPEAT       1     -2.72764388        1.03492496          -2.636             0.0086
GRADE56      1      2.28740760        0.65760631           3.478             0.0005
ABSMUCH      1     0.001592324        1.00072885           0.002             0.9987
EPOSTHS1     1      1.51162627        0.97882976           1.544             0.1230
FAMSIZE      1     -0.39928006        0.21070127          -1.895             0.0585
BRTHORDR     1    -0.001765108        0.18303764          -0.010             0.9923
LANGPROB     1     -2.56292454        1.77622483          -1.443             0.1495
MALE         1      3.33557171        1.20372858           2.771             0.0057
BLACK        1      1.40533680        2.04635845           0.687             0.4925
HISPANIC     1      0.65292580        2.38298576           0.274             0.7842
ASIAN        1     -0.24677942        2.62377083          -0.094             0.9251
BOSTON       1     -6.01010269        1.22643515          -4.900             0.0001
FRESNO       1     -7.28570389        1.28837263          -5.655             0.0001
SANDIEGO     1     -2.68354260        1.35462213          -1.981             0.0480
```

Table E.18

REGRESSION ANALYSIS OF MATH SCORES WITH SITE-TREATMENT INTERACTION TERMS (PORTLAND EXCLUDED)

```
DEP VARIABLE: MNCE
ANALYSIS OF VARIANCE

                        SUM OF          MEAN
SOURCE      DF         SQUARES        SQUARE         F VALUE          PROB>F

MODEL       21       182265.33    8679.30123          47.248          0.0001
ERROR      651       119586.94     183.69729
C TOTAL    672       301852.26

     ROOT MSE          13.5535      R-SQUARE          0.6038
     DEP MEAN         49.86716      ADJ R-SQ          0.5910
     C.V.              27.1792

PARAMETER ESTIMATES

                        PARAMETER          STANDARD         T FOR H0:
VARIABLE    DF           ESTIMATE             ERROR       PARAMETER=0         PROB > |T|

INTERCEP     1       -83.39519584        9.97021155            -8.364             0.0001
TREAT        1         0.78956661        2.08847285             0.378             0.7055
BOS_T        1        -0.19913269        2.86605203            -0.069             0.9446
FRESNO_T     1         1.71329734        2.83042174             0.605             0.5452
DIEGO_T      1         1.80930104        3.41676286             0.530             0.5966
STDMATH1     1         0.12295490        0.01000759            12.286             0.0001
ENJMATH1     1        -3.01702909        0.48692090            -6.196             0.0001
STDREAD1     1         0.05930823        0.01217512             4.871             0.0001
REPEAT       1        -3.13416287        1.32450659            -2.366             0.0183
GRADE56      1         1.27641438        0.84866280             1.504             0.1331
ABSMUCH      1        -1.01193912        1.29274561            -0.783             0.4340
EPOSTHS1     1         3.05013402        1.25068409             2.439             0.0150
FAMSIZE      1         0.48146369        0.27563487             1.747             0.0812
BRTHORDR     1       -0.008737224        0.22917345            -0.038             0.9696
LANGPROB     1         1.58778632        2.27343047             0.698             0.4852
MALE         1        -0.29240624        1.08708666            -0.269             0.7880
BLACK        1        -2.17203299        1.98832427            -1.092             0.2751
HISPANIC     1        -2.47242314        2.28860554            -1.080             0.2804
ASIAN        1        10.82561836        2.78671231             3.885             0.0001
BOSTON       1        -7.69447116        2.14302261            -3.590             0.0004
FRESNO       1        -5.29532329        2.14630007            -2.467             0.0139
SANDIEGO     1        -0.82632412        2.44258209            -0.338             0.7352
```

Table E.19

REGRESSION ANALYSIS OF MATH SCORES WITH RACIAL/ETHNIC- AND SEX-TREATMENT INTERACTION TERMS (PORTLAND EXCLUDED)

```
DEP VARIABLE: MNCE
ANALYSIS OF VARIANCE

                        SUM OF          MEAN
SOURCE      DF         SQUARES        SQUARE         F VALUE         PROB)F

MODEL       22       182412.99    8291.49975          45.123         0.0001
ERROR      650       119439.27     183.75272
C TOTAL    672       301852.26

     ROOT MSE         13.55554      R-SQUARE          0.6043
     DEP MEAN         49.86716      ADJ R-SQ          0.5909
     C.V.              27.1833

PARAMETER ESTIMATES

                         PARAMETER         STANDARD      T FOR HO:
VARIABLE     DF          ESTIMATE            ERROR    PARAMETER=0      PROB ) :T:

INTERCEP      1      -84.62674984       9.93213243         -8.521          0.0001
TREAT         1        5.53969862       3.94069077          1.406          0.1603
BLACK_T       1       -3.60010638       4.02232173         -0.895          0.3711
HISP_T        1       -5.27238253       4.49819258         -1.172          0.2416
ASIAN_T       1       -3.17238436       4.69614288         -0.676          0.4996
MALE_T        1       -0.84429578       2.12376710         -0.398          0.6911
STDMATH1      1        0.12358368       0.01000678         12.350          0.0001
ENJMATH1      1       -3.03211589       0.48728600         -6.222          0.0001
STDREAD1      1        0.05794017       0.01215076          4.768          0.0001
REPEAT        1       -3.17790338       1.32170454         -2.404          0.0165
GRADE56       1        1.29817624       0.84989707          1.527          0.1271
ABSMUCH       1       -0.92246017       1.29570899         -0.712          0.4768
EPOSTHS1      1        3.00368519       1.24959563          2.404          0.0165
FAMSIZE       1        0.47362748       0.27550692          1.719          0.0861
BRTHORDR      1      -0.008074222       0.22927979         -0.035          0.9719
LANGPROB      1        1.72307335       2.27416134          0.758          0.4489
MALE          1        0.04937037       1.53585657          0.032          0.9744
BLACK         1       -1.02095828       2.49317879         -0.410          0.6823
HISPANIC      1       -0.42399886       2.93088951         -0.145          0.8850
ASIAN         1       11.67764121       3.31396589          3.524          0.0005
BOSTON        1       -7.86218119       1.60409693         -4.901          0.0001
FRESNO        1       -4.46311978       1.66828066         -2.675          0.0077
SANDIEGO      1        0.05668269       1.76333483          0.032          0.9744
```

Table E.20

REGRESSION ANALYSIS OF READING SCORES (IN NCES) INCLUDING SAN DIEGO 1986 PRE-TEST SCORES WITH SITE-TREATMENT INTERACTION TERMS (PORTLAND EXCLUDED)

```
DEP VARIABLE: RNCE3
ANALYSIS OF VARIANCE

                     SUM OF          MEAN
SOURCE     DF       SQUARES        SQUARE        F VALUE       PROB)F

MODEL      20     287291.88   14364.59386         93.837       0.0001
ERROR     743     113738.41     153.07996
C TOTAL   763     401030.29

     ROOT MSE      12.37255      R-SQUARE         0.7164
     DEP MEAN      48.56152      ADJ R-SQ         0.7088
     C.V.          25.47809

PARAMETER ESTIMATES

                      PARAMETER          STANDARD      T FOR H0:
VARIABLE   DF          ESTIMATE             ERROR    PARAMETER=0       PROB ) !T!

INTERCEP    1        -118.69520        8.31677436        -14.272           0.0001
TREAT       1        3.34816113        1.89089606          1.771           0.0770
BOS_T       1       -2.23127096        2.59495969         -0.860           0.3902
FRESNO_T    1       -2.93270832        2.58241038         -1.136           0.2565
DIEGO_T     1      -10.09340471        2.62384264         -3.847           0.0001
STDREAD1    1        0.19888814       0.008657795         22.372           0.0001
ENJREAD1    1       -1.11441943        0.44800363         -2.488           0.0131
REPEAT      1       -1.93895592        1.13184548         -1.713           0.0871
GRADE56     1        2.25868881        0.63068854          3.581           0.0004
ABSMUCH     1        0.05293522        1.10048936          0.048           0.9616
EPOSTHS1    1        1.77496519        1.07704423          1.648           0.0998
FAMSIZE     1       -0.31596456        0.23469994         -1.346           0.1786
BRTHORDR    1        0.08145265        0.20318439          0.401           0.6886
LANGPROB    1       -2.67615072        1.86899227         -1.432           0.1526
MALE        1        2.46242645        0.92520743          2.661           0.0079
BLACK       1       -0.85580408        1.75325032         -0.488           0.6256
HISPANIC    1       -2.06377683        2.00184554         -1.031           0.3029
ASIAN       1        3.59899137        2.31660659          1.554           0.1207
BOSTON      1       -4.20890077        1.90691341         -2.207           0.0276
FRESNO      1       -5.06486116        1.94043875         -2.610           0.0092
SANDIEGO    1       31.29118499        1.79574076         17.425           0.0001
```

Table E.21

REGRESSION ANALYSIS OF READING SCORES (IN NCES) INCLUDING SAN DIEGO 1986 PRE-TEST SCORES WITH RACIAL/ETHNIC- AND SEX-TREATMENT INTERACTION TERMS
(PORTLAND EXCLUDED)

```
DEP VARIABLE: RNCE3
ANALYSIS OF VARIANCE

                        SUM OF          MEAN
SOURCE      DF         SQUARES        SQUARE       F VALUE       PROB>F

MODEL       21       285763.23   13607.77287        87.596       0.0001
ERROR      742       115267.06     155.34644
C TOTAL    763       401030.29

    ROOT MSE          12.46381      R-SQUARE        0.7126
    DEP MEAN          48.56152      ADJ R-SQ        0.7044
    C.V.              25.66601

PARAMETER ESTIMATES

                        PARAMETER          STANDARD     T FOR H0:
VARIABLE    DF           ESTIMATE             ERROR   PARAMETER=0      PROB > |T|

INTERCEP     1         -119.22862        8.43072312       -14.142          0.0001
TREAT        1         3.21979353        3.42621766         0.940          0.3477
BLACK_T      1        -4.13960891        3.49781321        -1.183          0.2370
HISP_T       1        -8.64002067        3.87355026        -2.231          0.0260
ASIAN_T      1        -3.89238309        3.96844958        -0.981          0.3270
MALE_T       1         1.79298452        1.83447560         0.977          0.3287
STDREAD1     1         0.20050482       0.008694425        23.061          0.0001
ENJREAD1     1        -1.19404038        0.45013478        -2.653          0.0082
REPEAT       1        -1.87339540        1.14137916        -1.641          0.1011
GRADE56      1         2.25860973        0.63770351         3.542          0.0004
ABSMUCH      1        -0.07005519        1.10859611        -0.063          0.9496
EPOSTHS1     1         1.61369078        1.08503294         1.487          0.1374
FAMSIZE      1        -0.31190077        0.23656457        -1.318          0.1878
BRTHORDR     1         0.07937065        0.20475471         0.388          0.6984
LANGPROB     1        -2.66040488        1.88759740        -1.409          0.1591
MALE         1         1.11102547        1.31109241         0.847          0.3970
BLACK        1         1.15690119        2.32725174         0.497          0.6193
HISPANIC     1         2.24463494        2.67002912         0.841          0.4008
ASIAN        1         5.49707455        2.86540527         1.918          0.0554
BOSTON       1        -5.12568966        1.41847277        -3.614          0.0003
FRESNO       1        -6.27291603        1.46641453        -4.278          0.0001
SANDIEGO     1        26.58841159        1.34091647        19.829          0.0001
```

Table E.22

REGRESSION ANALYSIS OF MATH SCORES (IN NCES) INCLUDING SAN DIEGO 1986 PRE-TEST SCORES WITH SITE-TREATMENT INTERACTION TERMS (PORTLAND EXCLUDED)

```
DEP VARIABLE: MNCE3
ANALYSIS OF VARIANCE

                         SUM OF          MEAN
SOURCE       DF         SQUARES        SQUARE        F VALUE        PROB>F

MODEL        21       362628.68   17268.03240         88.608        0.0001
ERROR       737       143627.66     194.88149
C TOTAL     758       506256.34

     ROOT MSE             13.96      R-SQUARE         0.7163
     DEP MEAN          56.77233      ADJ R-SQ         0.7082
     C.V.              24.58944

PARAMETER ESTIMATES

                          PARAMETER          STANDARD       T FOR H0:
VARIABLE     DF            ESTIMATE             ERROR     PARAMETER=0        PROB > |T|

INTERCEP      1        -89.30965913        9.36370789          -9.538            0.0001
TREAT         1          1.36480573        2.14797144           0.635            0.5254
BOS_T         1         -0.66252847        2.94826291          -0.225            0.8223
FRESNO_T      1          1.04863588        2.91158707           0.360            0.7188
DIEGO_T       1         -3.98832161        2.97236750          -1.342            0.1801
STDMATH1      1          0.12274817       0.009666386          12.698            0.0001
ENJMATH1      1         -2.99588450        0.47466109          -6.312            0.0001
STDREAD1      1          0.04762318        0.01178016           4.043            0.0001
REPEAT        1         -2.17277090        1.27944441          -1.698            0.0899
GRADE56       1          3.19376319        0.71411177           4.472            0.0001
ABSMUCH       1         -1.76263926        1.25457992          -1.405            0.1605
EPOSTHS1      1          2.18283198        1.21762229           1.793            0.0734
FAMSIZE       1          0.27860343        0.26944206           1.034            0.3015
BRTHORDR      1         -0.12506376        0.22435142          -0.557            0.5774
LANGPROB      1          0.39198674        2.11151437           0.186            0.8528
MALE          1          0.18893766        1.04667156           0.181            0.8568
BLACK         1         -2.56382637        1.96462745          -1.305            0.1923
HISPANIC      1         -0.55002130        2.24188418          -0.246            0.8060
ASIAN         1          5.89474130        2.65368931           2.221            0.0266
BOSTON        1         -9.00832262        2.18183041          -4.129            0.0001
FRESNO        1         -8.01431116        2.17525459          -3.684            0.0002
SANDIEGO      1         23.75814788        2.07327951          11.459            0.0001
```

Table E.23

REGRESSION ANALYSIS OF MATH SCORES (IN NCES) INCLUDING SAN DIEGO 1986 PRE-TEST SCORES WITH RACIAL/ETHNIC- AND SEX-TREATMENT INTERACTION TERMS (PORTLAND EXCLUDED)

```
DEP VARIABLE: MNCE3
ANALYSIS OF VARIANCE

                      SUM OF          MEAN
SOURCE      DF       SQUARES        SQUARE        F VALUE        PROB>F

MODEL       22     362397.13   16472.59687         84.276        0.0001
ERROR      736     143859.21     195.46088
C TOTAL    758     506256.34

     ROOT MSE       13.98073      R-SQUARE         0.7158
     DEP MEAN       56.77233      ADJ R-SQ         0.7073
     C.V.           24.62596

PARAMETER ESTIMATES

                     PARAMETER          STANDARD     T FOR H0:
VARIABLE   DF         ESTIMATE             ERROR   PARAMETER=0       PROB > |T|

INTERCEP    1     -90.72155994        9.35738901         -9.695          0.0001
TREAT       1       5.49935665        3.90888929          1.407          0.1599
BLACK_T     1      -4.80056753        3.98949983         -1.203          0.2292
HISP_T      1      -5.67544039        4.41314417         -1.286          0.1988
ASIAN_T     1      -3.52074193        4.52270595         -0.778          0.4365
MALE_T      1      -1.22218872        2.06109169         -0.593          0.5534
STDMATH1    1       0.12370231       0.009689529         12.767          0.0001
ENJMATH1    1      -2.99794850        0.47584139         -6.300          0.0001
STDREAD1    1       0.04694336        0.01176143          3.991          0.0001
REPEAT      1      -2.20707627        1.27966383         -1.725          0.0850
GRADE56     1       3.16994756        0.71781958          4.416          0.0001
ABSMUCH     1      -1.82534059        1.25698836         -1.452          0.1469
EPOSTHS1    1       2.12171428        1.22036936          1.739          0.0825
FAMSIZE     1       0.29245723        0.26966813          1.085          0.2785
BRTHORDR    1      -0.11640928        0.22481859         -0.518          0.6048
LANGPROB    1       0.43577873        2.11826271          0.206          0.8371
MALE        1       0.56157297        1.46202990          0.384          0.7010
BLACK       1      -0.66670960        2.48340423         -0.268          0.7884
HISPANIC    1       1.70852656        2.87548070          0.594          0.5526
ASIAN       1       7.12426434        3.15289280          2.260          0.0241
BOSTON      1      -9.34386726        1.62204319         -5.761          0.0001
FRESNO      1      -7.39577067        1.66764714         -4.435          0.0001
SANDIEGO    1      21.92102743        1.55168460         14.127          0.0001
```

Table E.24

REGRESSION ANALYSIS OF THE PERCENTAGE OF THE YEAR ABSENT WITH SITE-TREATMENT INTERACTION TERMS (SAN DIEGO EXCLUDED)

```
DEP VARIABLE: PCTABS
ANALYSIS OF VARIANCE

                        SUM OF            MEAN
SOURCE        DF       SQUARES          SQUARE        F VALUE        PROB)F

MODEL         22   20172.11796       916.91445          6.558        0.0001
ERROR        905     126535.68       139.81843
C TOTAL      927     146707.80

    ROOT MSE          11.82448      R-SQUARE           0.1375
    DEP MEAN          12.74887      ADJ R-SQ           0.1165
    C.V.              92.74927

PARAMETER ESTIMATES

                      PARAMETER            STANDARD       T FOR HO:
VARIABLE    DF         ESTIMATE               ERROR     PARAMETER=0        PROB > |T|

INTERCEP     1      18.17594962          7.48149462           2.429            0.0153
TREAT        1       0.27370129          1.62413077           0.169            0.8662
BOS_T        1      -3.06532759          2.23277300          -1.373            0.1701
FRESNO_T     1       0.72487482          2.21972084           0.327            0.7441
PORT_T       1      -1.74204225          2.29352451          -0.760            0.4477
STDREAD1     1     -0.006632948         0.009197038          -0.721            0.4710
STDMATH1     1     -0.009002181         0.007214256          -1.248            0.2124
ENJREAD1     1       0.43092160          0.39997732           1.077            0.2816
ENJMATH1     1      -0.12330293          0.36698734          -0.336            0.7370
REPEAT       1       1.22099599          0.97513129           1.252            0.2108
GRADE56      1       0.09497633          0.63607388           0.149            0.8813
ABSMUCH      1       6.09055634          0.89459717           6.808            0.0001
EPOSTHS1     1      -0.37990530          0.90310744          -0.421            0.6741
FAMSIZE      1      -0.39186560          0.19685380          -1.991            0.0468
BRTHORDR     1       0.16079038          0.17211616           0.934            0.3505
LANGPROB     1      -3.19613286          1.89325063          -1.688            0.0917
MALE         1      -1.29838810          0.79950612          -1.624            0.1047
BLACK        1       1.05968122          1.24294642           0.853            0.3941
HISPANIC     1       3.62186532          1.61349165           2.245            0.0250
ASIAN        1      -1.90272250          2.05638002          -0.925            0.3551
BOSTON       1       7.86956340          1.66283609           4.733            0.0001
FRESNO       1       1.27536259          1.72366537           0.740            0.4595
PORTLAND     1       5.21722451          1.67357779           3.117            0.0019
```

Table E.25

REGRESSION ANALYSIS OF THE PERCENTAGE OF THE YEAR ABSENT WITH RACIAL/ETHNIC- AND SEX-TREATMENT INTERACTION TERMS (SAN DIEGO EXCLUDED)

```
DEP VARIABLE: PCTABS
ANALYSIS OF VARIANCE

                      SUM OF          MEAN
SOURCE      DF       SQUARES        SQUARE        F VALUE        PROB)F

MODEL       23   19857.71510     863.37892          6.153        0.0001
ERROR      904     126850.08     140.32089
C TOTAL    927     146707.80

    ROOT MSE        11.84571      R-SQUARE         0.1354
    DEP MEAN        12.74887      ADJ R-SQ         0.1134
    C.V.            92.91577

PARAMETER ESTIMATES

                     PARAMETER          STANDARD       T FOR HO:
VARIABLE    DF        ESTIMATE             ERROR     PARAMETER=0       PROB ) :T:

INTERCEP     1     18.22595303        7.48400361           2.435           0.0151
TREAT        1     -0.50573664        2.36628560          -0.214           0.8308
BLACK_T      1     -0.11379999        2.42632454          -0.047           0.9626
HISP_T       1     -2.16842484        2.95883922          -0.733           0.4638
ASIAN_T      1      1.46796474        3.38613158           0.434           0.6647
MALE_T       1      0.09019143        1.57845391           0.057           0.9544
STDREAD1     1    -0.007911797       0.009207834          -0.859           0.3904
STDMATH1     1    -0.007881763       0.007235002          -1.089           0.2763
ENJREAD1     1      0.48228376        0.39956178           1.207           0.2277
ENJMATH1     1     -0.12641415        0.36709014          -0.344           0.7306
REPEAT       1      1.19731890        0.97734972           1.225           0.2209
GRADE56      1      0.13637142        0.63646476           0.214           0.8304
ABSMUCH      1      6.04153775        0.89652315           6.739           0.0001
EPOSTHS1     1     -0.36437386        0.90248746          -0.404           0.6865
FAMSIZE      1     -0.36208158        0.19683583          -1.840           0.0662
BRTHORDR     1      0.16283142        0.17248174           0.944           0.3454
LANGPROB     1     -3.11130219        1.90825725          -1.630           0.1034
MALE         1     -1.40061532        1.12901411          -1.241           0.2151
BLACK        1      0.97444404        1.64998852           0.591           0.5550
HISPANIC     1      4.45807031        2.10600695           2.117           0.0345
ASIAN        1     -2.92017873        2.54315475          -1.148           0.2512
BOSTON       1      6.35020375        1.23816807           5.129           0.0001
FRESNO       1      1.63168835        1.33032163           1.227           0.2203
PORTLAND     1      4.33006667        1.20539610           3.592           0.0003
```

Table E.26

REGRESSION ANALYSIS OF THE PERCENTAGE OF HIGH SCHOOL CREDITS EARNED WITH SITE-TREATMENT INTERACTION TERMS (ALL SITES INCLUDED)

```
DEP VARIABLE: PCTCRED
ANALYSIS OF VARIANCE

                    SUM OF           MEAN
SOURCE     DF      SQUARES         SQUARE        F VALUE        PROB>F

MODEL      39    249984.10     6409.84870         49.883        0.0001
ERROR     963    123743.97      128.49841
C TOTAL  1002    373728.07

    ROOT MSE      11.33571       R-SQUARE         0.6689
    DEP MEAN      24.62414       ADJ R-SQ         0.6555
    C.V.          46.03496

PARAMETER ESTIMATES

                   PARAMETER         STANDARD       T FOR H0:
VARIABLE   DF       ESTIMATE            ERROR     PARAMETER=0       PROB > |T|

INTERCEP    1     -236.90707      10.22928221         -23.160           0.0001
TREAT       1     2.38356522       1.69057843           1.410           0.1589
BOS_T       1    -1.64421049       2.30333218          -0.714           0.4755
FRESNO_T    1    -3.20289097       2.23676148          -1.432           0.1525
PORT_T      1    -3.62842301       2.36716069          -1.533           0.1256
DIEGO_T     1    -2.29807768       2.48235489          -0.926           0.3548
BOSTON      1    -3.49918357       1.76730595          -1.980           0.0480
FRESNO      1      236.30358      14.80084052          15.966           0.0001
PORTLAND    1      240.72734      14.72596544          16.347           0.0001
SANDIEGO    1     3.34180845       1.82562080           1.831           0.0675
BLACK_DB    1    -4.21305827       1.79121640          -2.352           0.0189
HISP_DB     1    -4.15755295       2.28914978          -1.816           0.0697
ASIAN_DB    1    -2.51170977       2.42596835          -1.035           0.3008
BLACK_DF    1     4.47258347       1.85524885           2.411           0.0161
HISP_DF     1     5.55705938       2.14527944           2.590           0.0097
ASIAN_DF    1     9.65345853       3.37125327           2.863           0.0043
BSTDR       1   -0.003972126       0.01139964          -0.348           0.7276
BSTDM       1    0.009744441      0.009276964           1.050           0.2938
BENJR       1    -0.49194110       0.47310709          -1.040           0.2987
BENJM       1     0.28337530       0.44888027           0.631           0.5280
BREPE       1    -0.37877879       1.20310516          -0.315           0.7530
BGR56       1    28.18084133       0.86555926          32.558           0.0001
BABSM       1    -5.08245907       1.22268396          -4.157           0.0001
BEHS1       1     1.81358300       1.10594192           1.640           0.1014
BFAMS       1     0.41490852       0.24620692           1.685           0.0923
BBRTH       1     0.05172566       0.23263906           0.222           0.8241
BLANP       1     0.30108503       1.97384989           0.153           0.8788
BMALE       1     0.24098487       0.99386337           0.242           0.8085
FSTDR       1    0.007042773       0.01273643           0.553           0.5804
FSTDM       1     0.01490158      0.009845017           1.514           0.1305
FENJR       1     0.49257738       0.59152537           0.833           0.4052
FENJM       1    -0.42637913       0.53951265          -0.790           0.4295
FREPE       1    -2.84149702       1.44474738          -1.967           0.0495
FGR56       1     0.04740743       0.88083617           0.054           0.9571
FABSM       1    -5.91683864       1.20262697          -4.920           0.0001
FEHS1       1     2.15697493       1.30468673           1.653           0.0986
FFAMS       1    -0.18355933       0.26618032          -0.690           0.4906
FBRTH       1    -0.28033063       0.21449915          -1.307           0.1916
FLANP       1    -0.57666454       2.85535498          -0.202           0.8400
FMALE       1    -0.43241934       1.12422631          -0.385           0.7006
```

Table E.27

REGRESSION ANALYSIS OF THE PERCENTAGE OF HIGH SCHOOL CREDITS WITH RACIAL/ETHNIC- AND SEX-TREATMENT INTERACTION TERMS (ALL SITES INCLUDED)

```
DEP VARIABLE: PCTCRED
ANALYSIS OF VARIANCE

                      SUM OF          MEAN
SOURCE      DF       SQUARES        SQUARE       F VALUE       PROB>F

MODEL       39     249861.61    6406.70783        49.809       0.0001
ERROR      963     123866.47     128.62561
C TOTAL   1002     373728.07

     ROOT MSE       11.34132      R-SQUARE        0.6686
     DEP MEAN       24.62414      ADJ R-SQ        0.6551
     C.V.           46.05774

PARAMETER ESTIMATES

                       PARAMETER        STANDARD      T FOR H0:
VARIABLE   DF           ESTIMATE           ERROR    PARAMETER=0     PROB > |T|

INTERCEP    1         -234.98871     10.15038341        -23.151         0.0001
TREAT       1         1.83565658      2.51127517          0.731         0.4650
BLACK_T     1        -1.43744260      2.55295018         -0.563         0.5735
HISP_T      1         0.35416724      2.94621391          0.120         0.9043
ASIAN_T     1        -0.44188179      3.10933534         -0.142         0.8870
MALE_T      1        -1.59927625      1.45853487         -1.096         0.2731
BOSTON      1        -4.41169195      1.30951633         -3.369         0.0008
FRESNO      1         234.13042      14.70090808         15.926         0.0001
PORTLAND    1         238.33533      14.59713821         16.328         0.0001
SANDIEGO    1         2.24916283      1.36684538          1.646         0.1002
BLACK_DB    1        -3.65173745      2.06218183         -1.771         0.0769
HISP_DB     1        -4.48297751      2.58061939         -1.737         0.0827
ASIAN_DB    1        -2.37925934      2.71993657         -0.875         0.3819
BLACK_DF    1         4.85852570      2.12536986          2.286         0.0225
HISP_DF     1         5.00750820      2.50773030          1.997         0.0461
ASIAN_DF    1         9.93680923      3.63630430          2.733         0.0064
BSTDR       1     -0.002698736      0.01139566         -0.237         0.8128
BSTDM       1      0.008467813     0.009298559          0.911         0.3627
BENJR       1        -0.46871763      0.47163476         -0.994         0.3206
BENJM       1         0.24175145      0.44821618          0.539         0.5898
BREPE       1        -0.45494230      1.20764076         -0.377         0.7065
BGR56       1        28.02337490      0.86117967         32.541         0.0001
BABSM       1        -5.10563404      1.22322456         -4.174         0.0001
BEHS1       1         1.95013838      1.10302210          1.768         0.0774
BFAMS       1         0.40737879      0.24547977          1.660         0.0973
BBRTH       1         0.06065533      0.23246930          0.261         0.7942
BLANP       1         0.40010011      1.98277027          0.202         0.8401
BMALE       1         0.94635061      1.22513817          0.772         0.4400
FSTDR       1       0.006528893      0.01280524          0.510         0.6103
FSTDM       1         0.01497636     0.009859035          1.519         0.1291
FENJR       1         0.48003763      0.59210553          0.811         0.4177
FENJM       1        -0.39585358      0.53986726         -0.733         0.4636
FREPE       1        -2.82640856      1.44542881         -1.955         0.0508
FGR56       1       0.005373177      0.88258775          0.006         0.9951
FABSM       1        -5.93670880      1.20620936         -4.922         0.0001
FEHS1       1         2.16971430      1.30607567          1.661         0.0970
FFAMS       1        -0.19184586      0.26518177         -0.723         0.4696
FBRTH       1        -0.29726815      0.21483448         -1.384         0.1668
FLANP       1        -1.02349160      2.86090145         -0.358         0.7206
FMALE       1         0.39975805      1.34366014          0.298         0.7661
```

Table E.28

LOGIT ANALYSIS OF THE PROBABILITY OF PROMOTION
(ALL SITES INCLUDED)

```
                     DEPENDENT VARIABLE: PROMOTED

          1255 OBSERVATIONS
           296  PROMOTED=  0
           959  PROMOTED=  1
           321 OBSERVATIONS DELETED DUE TO MISSING VALUES

VARIABLE          BETA          STD. ERROR       CHI-SQUARE     P        R

INTERCEPT   -0.51030125      1.36392154         0.14    0.7083
TREAT        0.29245286      0.14343245         4.16    0.0415    0.040
STDREAD1     0.00262271      0.00167620         2.45    0.1177    0.018
STDMATH1     0.00086278      0.00128254         0.45    0.5011    0.000
ENJREAD1     0.12515369      0.07249742         2.98    0.0843    0.027
ENJMATH1    -0.11616099      0.06619638         3.08    0.0793   -0.028
REPEAT      -0.44466104      0.16842160         6.97    0.0083   -0.060
GRADE56     -0.09259334      0.11544209         0.64    0.4225    0.000
ABSMUCH     -0.84950246      0.15284567        30.89    0.0000   -0.145
EPOSTHS1     0.43438674      0.15663851         7.69    0.0056    0.064
FAMSIZE      0.02402417      0.03660262         0.43    0.5116    0.000
BRTHORDR    -0.03800115      0.03116882         1.49    0.2228    0.000
LANGPROB    -0.31440226      0.35420976         0.79    0.3747    0.000
MALE        -0.41944986      0.14694799         8.15    0.0043   -0.067
BLACK        0.09583599      0.22222222         0.19    0.6663    0.000
HISPANIC    -0.03705485      0.27974889         0.02    0.8946    0.000
ASIAN        0.75402068      0.40362787         3.49    0.0617    0.033
BOSTON      -0.29867193      0.22824457         1.71    0.1907    0.000
FRESNO      -0.04153509      0.23980000         0.03    0.8625    0.000
PORTLAND     0.40567688      0.23692337         2.93    0.0868    0.026
SANDIEGO     1.05711993      0.30244110        12.22    0.0005    0.086
```

Table E.29

LOGIT ANALYSIS OF THE PROBABILITY OF PROMOTION WITH SITE-TREATMENT INTERACTION TERMS (ALL SITES INCLUDED)

DEPENDENT VARIABLE: PROMOTED

1255 OBSERVATIONS
296 PROMOTED= 0
959 PROMOTED= 1
321 OBSERVATIONS DELETED DUE TO MISSING VALUES

VARIABLE	BETA	STD. ERROR	CHI-SQUARE	P	R
INTERCEPT	-0.41481911	1.37770700	0.09	0.7633	
TREAT	0.09608263	0.30757736	0.10	0.7547	0.000
STDREAD1	0.00261971	0.00168182	2.43	0.1193	0.018
STDMATH1	0.00088179	0.00128534	0.47	0.4927	0.000
ENJREAD1	0.13014413	0.07284535	3.19	0.0740	0.029
ENJMATH1	-0.11769335	0.06637383	3.14	0.0762	-0.029
REPEAT	-0.45430059	0.16894022	7.23	0.0072	-0.062
GRADE56	-0.09502970	0.11588734	0.67	0.4122	0.000
ABSMUCH	-0.85213278	0.15317715	30.95	0.0000	-0.145
EPOSTHS1	0.44387544	0.15714436	7.98	0.0047	0.066
FAMSIZE	0.02486695	0.03680176	0.46	0.4992	0.000
BRTHORDR	-0.03827537	0.03120284	1.50	0.2199	0.000
LANGPROB	-0.30293313	0.35498457	0.73	0.3935	0.000
MALE	-0.41777605	0.14731111	8.04	0.0046	-0.066
BLACK	0.08693832	0.22270280	0.15	0.6963	0.000
HISPANIC	-0.05418729	0.28079929	0.04	0.8470	0.000
ASIAN	0.73291276	0.40473641	3.28	0.0702	0.031
BOSTON	-0.46967804	0.30119228	2.43	0.1189	-0.018
FRESNO	-0.09437188	0.30737614	0.09	0.7588	0.000
PORTLAND	0.25966503	0.31922401	0.66	0.4160	0.000
SANDIEGO	0.97693011	0.39685332	6.06	0.0138	0.054
BOS_T	0.36587649	0.42020361	0.76	0.3839	0.000
FRESNO_T	0.12370752	0.41623984	0.09	0.7663	0.000
PORT_T	0.31098507	0.45612105	0.46	0.4954	0.000
DIEGO_T	0.17421982	0.58257769	0.09	0.7649	0.000

Table E.30

LOGIT ANALYSIS OF THE PROBABILITY OF PROMOTION WITH RACIAL/ETHNIC- AND SEX-TREATMENT INTERACTION TERMS (ALL SITES INCLUDED)

DEPENDENT VARIABLE: PROMOTED

1255 OBSERVATIONS
296 PROMOTED= 0
959 PROMOTED= 1
321 OBSERVATIONS DELETED DUE TO MISSING VALUES

VARIABLE	BETA	STD. ERROR	CHI-SQUARE	P	R
INTERCEPT	-0.28467444	1.38072658	0.04	0.8367	
TREAT	-0.19931578	0.42775228	0.22	0.6412	0.000
STDREAD1	0.00279602	0.00169660	2.72	0.0993	0.023
STDMATH1	0.00071549	0.00129015	0.31	0.5792	0.000
ENJREAD1	0.12542194	0.07274227	2.97	0.0847	0.027
ENJMATH1	-0.11788049	0.06637737	3.15	0.0757	-0.029
REPEAT	-0.44655239	0.16935109	6.95	0.0084	-0.060
GRADE56	-0.09486852	0.11633137	0.67	0.4148	0.000
ABSMUCH	-0.83024281	0.15355934	29.23	0.0000	-0.141
EPOSTHS1	0.44253042	0.15724033	7.92	0.0049	0.066
FAMSIZE	0.02371441	0.03676406	0.42	0.5189	0.000
BRTHORDR	-0.03975685	0.03128592	1.61	0.2038	0.000
LANGPROB	-0.28141412	0.35536747	0.63	0.4284	0.000
MALE	-0.26412635	0.19958110	1.75	0.1857	0.000
BLACK	-0.27572329	0.30526657	0.82	0.3664	0.000
HISPANIC	-0.43286742	0.37051182	1.36	0.2427	0.000
ASIAN	0.55461640	0.50585548	1.20	0.2729	0.000
BLACK_T	0.78501984	0.43400659	3.27	0.0705	0.030
HISP_T	0.81540118	0.51508190	2.51	0.1134	0.019
ASIAN_T	0.39755451	0.66117457	0.36	0.5477	0.000
MALE_T	-0.29292547	0.29091889	1.01	0.3140	0.000
BOSTON	-0.31512465	0.22893463	1.89	0.1687	0.000
FRESNO	-0.04624829	0.24027149	0.04	0.8474	0.000
PORTLAND	0.39986668	0.23743514	2.84	0.0922	0.025
SANDIEGO	1.06763316	0.30302594	12.41	0.0004	0.087

Table E.31

LOGIT ANALYSIS OF THE RESTRICTIVE MEASURE OF DROPOUTS (ALL SITES INCLUDED)

DEPENDENT VARIABLE: DROPOUT

1027 OBSERVATIONS
960 DROPOUT = 0
67 DROPOUT = 1
247 OBSERVATIONS DELETED DUE TO MISSING VALUES

VARIABLE	BETA	STD. ERROR	CHI-SQUARE	P	R
INTERCEPT	-1.93845176	2.30860971	0.71	0.4011	
TREAT	-0.09378350	0.26800207	0.12	0.7264	0.000
STDREAD1	-0.00047711	0.00314362	0.02	0.8794	0.000
STDMATH1	-0.00314726	0.00259994	1.47	0.2261	0.000
ENJREAD1	0.27961964	0.14033988	3.97	0.0463	0.063
ENJMATH1	0.09193573	0.12817932	0.51	0.4732	0.000
REPEAT	0.07787436	0.30799272	0.06	0.8004	0.000
GRADE56	0.08817634	0.14335096	0.38	0.5385	0.000
ABSMUCH	0.46982713	0.29816414	2.48	0.1151	0.031
EPOSTHS1	-0.30720774	0.29085272	1.12	0.2909	0.000
FAMSIZE	-0.11660064	0.07770566	2.25	0.1335	-0.023
BRTHORDR	0.04900568	0.05446559	0.81	0.3683	0.000
LANGPROB	-0.53602472	0.62317230	0.74	0.3897	0.000
MALE	0.19165030	0.27234371	0.50	0.4816	0.000
BLACK	-0.35201892	0.43871267	0.64	0.4223	0.000
HISPANIC	0.18990597	0.50110438	0.14	0.7047	0.000
ASIAN	-0.60098958	0.68677294	0.77	0.3815	0.000
BOSTON	0.27987365	0.45580218	0.38	0.5392	0.000
FRESNO	-0.57095796	0.53123091	1.16	0.2825	0.000
SANDIEGO	1.83521856	0.43183190	18.06	0.0000	0.180

Table E.32

LOGIT ANALYSIS OF THE RESTRICTIVE MEASURE OF DROPOUTS WITH SITE-TREATMENT INTERACTION TERMS (ALL SITES INCLUDED)

DEPENDENT VARIABLE: DROPOUT

1027 OBSERVATIONS
960 DROPOUT = 0
67 DROPOUT = 1
247 OBSERVATIONS DELETED DUE TO MISSING VALUES

VARIABLE	BETA	STD. ERROR	CHI-SQUARE	P	R
INTERCEPT	-1.62276498	2.33551308	0.48	0.4872	
TREAT	-0.39989222	0.66967525	0.36	0.5504	0.000
STDREAD1	-0.00044197	0.00314370	0.02	0.8882	0.000
STDMATH1	-0.00336263	0.00261279	1.66	0.1981	0.000
ENJREAD1	0.28152447	0.13989699	4.05	0.0442	0.064
ENJMATH1	0.08696216	0.12786821	0.46	0.4964	0.000
REPEAT	0.07801152	0.30910100	0.06	0.8007	0.000
GRADE56	0.08266706	0.13980330	0.35	0.5543	0.000
ABSMUCH	0.48089274	0.29847148	2.60	0.1071	0.035
EPOSTHS1	-0.28312019	0.29193009	0.94	0.3321	0.000
FAMSIZE	-0.12333838	0.07770414	2.52	0.1124	-0.032
BRTHORDR	0.05057950	0.05451670	0.86	0.3535	0.000
LANGPROB	-0.54578922	0.62149889	0.77	0.3798	0.000
MALE	0.21066871	0.27311203	0.60	0.4405	0.000
BLACK	-0.36861282	0.43994015	0.70	0.4021	0.000
HISPANIC	0.17963055	0.50238067	0.13	0.7207	0.000
ASIAN	-0.60836896	0.68561350	0.79	0.3749	0.000
BOSTON	0.04232109	0.60312938	0.00	0.9441	0.000
FRESNO	-1.07885851	0.75273457	2.05	0.1518	-0.010
SANDIEGO	1.82263015	0.54296425	11.27	0.0008	0.137
BOS_T	0.52894613	0.86385977	0.37	0.5403	0.000
FRESNO_T	1.01414984	1.00674167	1.01	0.3138	0.000
DIEGO_T	0.08752151	0.77237625	0.01	0.9098	0.000

Table E.33

LOGIT ANALYSIS OF THE RESTRICTIVE MEASURE OF DROPOUTS WITH RACIAL/ETHNIC- AND SEX-TREATMENT INTERACTION TERMS (ALL SITES INCLUDED)

DEPENDENT VARIABLE: DROPOUT

1027 OBSERVATIONS
960 DROPOUT = 0
67 DROPOUT = 1
247 OBSERVATIONS DELETED DUE TO MISSING VALUES

VARIABLE	BETA	STD. ERROR	CHI-SQUARE	P	R
INTERCEPT	-1.55059211	2.31519628	0.45	0.5030	
TREAT	-0.79144719	0.67450634	1.38	0.2406	0.000
STDREAD1	-0.00034133	0.00314194	0.01	0.9135	0.000
STDMATH1	-0.00324977	0.00261839	1.54	0.2146	0.000
ENJREAD1	0.28181855	0.14114329	3.99	0.0459	0.063
ENJMATH1	0.09206679	0.12819674	0.52	0.4727	0.000
REPEAT	0.08353816	0.30845645	0.07	0.7865	0.000
GRADE56	0.07347302	0.14390983	0.26	0.6097	0.000
ABSMUCH	0.45652652	0.29898949	2.33	0.1268	0.026
EPOSTHS1	-0.27222676	0.29264799	0.87	0.3523	0.000
FAMSIZE	-0.11782939	0.07761990	2.30	0.1290	-0.025
BRTHORDR	0.04577825	0.05482153	0.70	0.4037	0.000
LANGPROB	-0.47812702	0.62549412	0.58	0.4446	0.000
MALE	-0.00654230	0.37496621	0.00	0.9861	0.000
BLACK	-0.49197783	0.52742965	0.87	0.3509	0.000
HISPANIC	-0.23769728	0.62106025	0.15	0.7019	0.000
ASIAN	-0.64043705	0.69194720	0.86	0.3547	0.000
BLACK_T	0.36028630	0.69680648	0.27	0.6051	0.000
HISP_T	0.99234117	0.80048206	1.54	0.2151	0.000
MALE_T	0.46178331	0.53809609	0.74	0.3908	0.000
BOSTON	0.27260094	0.45735398	0.36	0.5511	0.000
FRESNO	-0.57944459	0.53325795	1.18	0.2772	0.000
SANDIEGO	1.82411925	0.43330491	17.72	0.0000	0.178

Table E.34

LOGIT ANALYSIS OF THE LESS RESTRICTIVE MEASURE OF DROPOUTS (ALL SITES INCLUDED)

DEPENDENT VARIABLE: DROP2

1276 OBSERVATIONS
1077 DROP2 = 0
199 DROP2 = 1
300 OBSERVATIONS DELETED DUE TO MISSING VALUES

VARIABLE	BETA	STD. ERROR	CHI-SQUARE	P	R
INTERCEPT	-0.69294446	1.54533725	0.20	0.6539	
TREAT	0.04919697	0.16568047	0.09	0.7665	0.000
STDREAD1	-0.00516205	0.00196652	6.89	0.0087	-0.067
STDMATH1	-0.00141086	0.00146652	0.93	0.3360	0.000
ENJREAD1	0.08552943	0.08658335	0.98	0.3232	0.000
ENJMATH1	0.14017645	0.07888653	3.16	0.0756	0.032
REPEAT	0.36331715	0.19776979	3.37	0.0662	0.035
GRADE56	0.15629592	0.11564273	1.83	0.1765	0.000
ABSMUCH	0.37314909	0.18004186	4.30	0.0382	0.046
EPOSTHS1	-0.03839493	0.18980928	0.04	0.8397	0.000
FAMSIZE	-0.02789179	0.04275000	0.43	0.5141	0.000
BRTHORDR	0.02668166	0.03384498	0.62	0.4305	0.000
LANGPROB	0.07079410	0.38957465	0.03	0.8558	0.000
MALE	0.30381555	0.17044136	3.18	0.0747	0.033
BLACK	-0.32446204	0.26553020	1.49	0.2217	0.000
HISPANIC	-0.27688929	0.31763186	0.76	0.3834	0.000
ASIAN	-0.52114697	0.45265332	1.33	0.2496	0.000
BOSTON	0.33023271	0.42637423	0.60	0.4386	0.000
FRESNO	2.32044335	0.36746619	39.88	0.0000	0.185
PORTLAND	2.01290926	0.35845669	31.53	0.0000	0.164
SANDIEGO	1.75147137	0.38267610	20.95	0.0000	0.131

Table E.35

LOGIT ANALYSIS OF THE LESS RESTRICTIVE MEASURE OF DROPOUTS WITH SITE-TREATMENT INTERACTION TERMS (ALL SITES INCLUDED)

DEPENDENT VARIABLE: DROP2

1276 OBSERVATIONS
1077 DROP2 = 0
199 DROP2 = 1
300 OBSERVATIONS DELETED DUE TO MISSING VALUES

VARIABLE	BETA	STD. ERROR	CHI-SQUARE	P	R
INTERCEPT	-0.40834725	1.56311892	0.07	0.7939	
TREAT	-0.57212120	0.64603474	0.78	0.3758	0.000
STDREAD1	-0.00506646	0.00197177	6.60	0.0102	-0.065
STDMATH1	-0.00153054	0.00147448	1.08	0.2993	0.000
ENJREAD1	0.09536329	0.08683436	1.21	0.2721	0.000
ENJMATH1	0.13881397	0.07880818	3.10	0.0782	0.032
REPEAT	0.36748754	0.19807370	3.44	0.0636	0.036
GRADE56	0.15505326	0.11426278	1.84	0.1748	0.000
ABSMUCH	0.37318660	0.18053786	4.27	0.0387	0.045
EPOSTHS1	-0.01893596	0.19017043	0.01	0.9207	0.000
FAMSIZE	-0.03203992	0.04308740	0.55	0.4571	0.000
BRTHORDR	0.02813415	0.03391841	0.69	0.4068	0.000
LANGPROB	0.05965620	0.38856756	0.02	0.8780	0.000
MALE	0.30474262	0.17094348	3.18	0.0746	0.033
BLACK	-0.34026880	0.26604126	1.64	0.2009	0.000
HISPANIC	-0.29744624	0.31839391	0.87	0.3502	0.000
ASIAN	-0.53510051	0.45148899	1.40	0.2359	0.000
BOSTON	0.03431762	0.56635214	0.00	0.9517	0.000
FRESNO	2.00576316	0.45907275	19.09	0.0000	0.124
PORTLAND	1.62411207	0.46516095	12.19	0.0005	0.096
SANDIEGO	1.64095599	0.48425228	11.48	0.0007	0.093
BOS_T	0.66784943	0.84207596	0.63	0.4277	0.000
FRESNO_T	0.72089228	0.70372001	1.05	0.3056	0.000
PORT_T	0.84645168	0.71890176	1.39	0.2390	0.000
DIEGO_T	0.30364909	0.74862611	0.16	0.6850	0.000